A COMPARISON OF THE PREDICTIVE POWER OF CLASSROOM OBSERVATION SYSTEMS AND ASPY PROCESS SCALES WITH REGARD TO SELECTED EDUCATIONAL OUTCOMES

By

FLORA NELL ROEBUCK

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF EDUCATION

UNIVERSITY OF FLORIDA

1975

ACKNOWLEDGMENTS

The acknowledgment of the debt I owe to so many of the faculty of the University of Florida is not an easy task as I find it difficult to pinpoint specific acts in what was a continuous process of inspiration and growth. But there was the special inspiration I found in Dr. A. W. Combs whose teaching and whose publications did so much to crystallize my experiencing and help me establish my commitment to the humanistic orientation in my profession and in my life. Dr. Don Avila never failed in his support and his exemplification of the impact that a caring person could have on the lives of those around him. Dr. Walter Busby was always there, providing informal learning situations in which the concepts from formal courses could be clarified and extended. And Dr. Franz Epting's delight and joy in theory and in intellectual problem-solving made learning a challenge rather than a threat. For their support, I give my thanks and my gratitude without stint, as their caring was without stint.

A special thank you is also extended to Dr. Robert Soar whose generosity with the data from his research project and whose help over the rough spots made this dissertation possible. To Dr. Dave Aspy, who developed the technology used in this dissertation and who remains a constant source of inspiration and of challenge, I also say a grateful "Thank you."

And last, as they were first, my love and my gratitude go to my parents, Mr. and Mrs. D. A. Roebuck, whose love and support have always made my tasks--from first grade math to this doctoral dissertation--possible of accomplishment.

GRANT

Part of the data used in this study were collected under National Institute of Mental Health Research Grant No. 15891-01 to the Institute for Development of Human Resources, University of Florida, Gainesville, Florida.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS ii

GRANT iv

LIST OF TABLES. ix

LIST OF FIGURES x

ABSTRACT. xi

CHAPTER

I INTRODUCTION. 1

- Origin of the Problem. 1
- Background of the Problem. 2
- Purpose of the Study 3
- Definition of Terms. 5
  - Educational Outcomes. 5
  - Specific-Behavior Measures. 6
  - Holistic Measures 6
  - More Highly Predictive. 10
- Rationale for the Study. 11

II REVIEW OF LITERATURE. 13

- Research on Teacher Effectiveness. 13
  - Holistic Variables. 14
  - Specific-Behavior Variables 17
  - Summary 18
- Background of the Selected Holistic Instrument--the Aspy Scales for Rating Interpersonal Processes 20
  - Theoretical Base. 20
  - Early Studies 21
  - Scales Reformulated 23
- Background of Selected Specific-Behavior Instruments--RCS and FLACCS Systematic Classroom Observation Systems. 24
  - Early Studies 25
  - Later Development 26
- Research Related to Study Assumptions. . . 28
  - Assumption 1. 29

TABLE OF CONTENTS (continued)

CHAPTER Page

II (cont.) Assumption 2. 30
Assumption 3. 31
Assumption 4. 32

III DESIGN OF THE STUDY 34

Overall Design of the Study. 34
Hypothesis 34
Statement of Null Hypotheses 35
Sample 36
Data Collection. 39
- Educational Outcome Indices 39
- I.Q. Test 40
- Florida Climate and Control System (FLACCS). 40
- Reciprocal Category System (RCS). 41
- Aspy Scales of Meaning, Genuineness, and Respect 42

Instrumentation. 44
- The Reciprocal Category System (RCS). . . 44
- Florida Climate and Control System (FLACCS). 45
- Aspy Rating Scales of Meaning, Genuineness, and Respect 47
- Children's Manifest Anxiety Scale 49
- Dependence-Proneness Scale. 50
- How I See Myself Scale. 50
- Iowa Tests of Basic Skills. 51
- Junior Index of Motivation. 52
- Lorge-Thorndike Intelligence Test 52
- Social Reaction Inventory 52
- Torrance Tests of Creative Thinking . . . 53

Statistical Procedures 54
- Procedures for Splitting the Sample . . . 55
- Calculation of Gain Scores. 56
- Procedures for Building and Using the Prediction Models 56
- Testing of Hypotheses 59

Bias of the Study. 60

IV RESULTS OF ANALYSIS 64

Means and Standard Deviations of Sub-Samples. 64
Results of Regression Analyses 64
Tests of Null Hypotheses 67

TABLE OF CONTENTS (continued)

CHAPTER Page

V DISCUSSION AND IMPLICATIONS 75

Discussion 75
Implications for Further Research. 78
True Effects of the Mediational Elements. 78
Differential Predictive Power of Holistic Model. 79
Comparison with Perceptual Variables. . 81
General Research Implications 82
Implications for Educational Practice. . . 87
Teacher Training. 87
Accountability. 89
Summary. 91

REFERENCES. 93

APPENDIX I Instruments for Measuring Classroom Process 104

Exhibit 1--ASPY RATING SCALES 105
Exhibit 2--FLORIDA CLIMATE AND CONTROL SYSTEM 108
Exhibit 3--SUMMARY OF CATEGORIES FOR THE RECIPROCAL CATEGORY SYSTEM . . 111

APPENDIX II Factor Structures 112

Exhibit 1--FACTOR STRUCTURE OF THE FLORIDA CLIMATE AND CONTROL SYSTEM . . 113
Exhibit 2--FACTOR STRUCTURE OF THE RECIPROCAL CATEGORY SYSTEM . . 117

APPENDIX III Regression Equations 125

Exhibit 1--REGRESSION EQUATIONS FROM ANALYSES USING HOLISTIC MEASURES AS PREDICTORS 126
Exhibit 2--REGRESSION EQUATIONS FROM ANALYSES USING SPECIFIC-BEHAVIOR MEASURES AS PREDICTORS. 129

TABLE OF CONTENTS (continued)

Page

APPENDIX IV Items from Scoring of Categorical Instruments 132

Exhibit 1--ITEMS SCORED FOR FLORIDA CLIMATE AND CONTROL SYSTEM . . 133
Exhibit 2--VARIABLES FOR RECIPROCAL CATEGORY SYSTEM. 135

BIOGRAPHICAL SKETCH 139

LIST OF TABLES

Table Page

1 DISTRIBUTION OF CLASSROOMS IN STUDY 36

2 STUDENT N WITH RACE/SEX DISTRIBUTION BY CLASSROOM 37

3 STUDENT DEMOGRAPHIC DATA. 38

4 INTER-RATER RELIABILITY FOR ASPY SCALES 49

5 PREDICTOR TERMS IN EACH OF THE REGRESSION MODELS. 58

6 MEANS AND STANDARD DEVIATIONS OF THE EDUCATIONAL OUTCOME INDICES FOR THE SAMPLE USED IN REGRESSION ANALYSES WITH HOLISTIC MEASURES AS PREDICTORS. 65

7 MEANS AND STANDARD DEVIATIONS OF THE EDUCATIONAL OUTCOME INDICES FOR THE SAMPLE USED IN REGRESSION ANALYSES WITH SPECIFIC-BEHAVIOR MEASURES AS PREDICTORS 66

8 MEANS AND STANDARD DEVIATIONS OF THE PREDICTOR TERMS IN THE REGRESSION ANALYSES USING HOLISTIC MEASURES 68

9 RESULTS OF REGRESSION ANALYSES USING HOLISTIC VARIABLES AS PREDICTORS. 69

10 MEANS AND STANDARD DEVIATIONS OF THE PREDICTOR TERMS IN REGRESSION ANALYSES USING SPECIFIC-BEHAVIOR MEASURES. 70

11 RESULTS OF REGRESSION ANALYSES USING SPECIFIC-BEHAVIOR VARIABLES AS PREDICTORS . . . 71

12 TESTS OF THE HYPOTHESES USING ONE-TAILED FISHER'S APPROXIMATION W STATISTIC. 73

LIST OF FIGURES

Figure Page

1. Teacher Effectiveness Paradigm. 2

2. Expanded Teacher Effectiveness Paradigm . . . 4

3. Response Surface for Change in Originality Factor of Creativity. 85

4. Response Surface for Change in Reading Achievement 86

Abstract of Dissertation Presented to the Graduate Council
of the University of Florida
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Education

A COMPARISON OF THE PREDICTIVE POWER OF CLASSROOM OBSERVATION SYSTEMS AND ASPY PROCESS SCALES WITH REGARD TO SELECTED EDUCATIONAL OUTCOMES

By

Flora Nell Roebuck

August, 1975

Chairman: Donald L. Avila
Major Department: Foundations of Education

This study compared the efficiency of holistic measures of classroom process against specific-behavior measures of classroom process in terms of how well they could predict student change on educational outcome indices. The holistic instrument used in the study was the Aspy Rating Scales of Meaning, Genuineness, and Respect. The specific-behavior instruments were the Reciprocal Category System and the Florida Climate and Control System. Educational outcome indices were 16 paper and pencil measures ranging from academic achievement tests through self-reports of self-concept.

The sample was sub-divided into two independent data sub-samples for the purpose of the statistical analysis. Each sub-sample provided the data set for regression analyses conducted with one of the two models. The multiple R's produced by regression analyses using holistic measures as predictor terms were compared against the multiple R's produced by regression analyses using specific-behavior measures as predictor terms. Null hypotheses were tested with the W statistic for Fisher's Approximation. It was found that the holistic measures did a significantly better job of predicting 3 of the outcome indices, performed as well for 12 of the outcome indices, and did not do as well as the specific-behavior measures for only 1 index. Thus, it was concluded that the study hypothesis of greater predictive power for the holistic measures was supported.

Supplementary analyses of the data indicated that the predictive power of the holistic measures increased when students were sampled from race-sex groups. Several specific studies arising from the current study were outlined and general research implications were stated. Implications for educational practice in the areas of teacher-training and accountability were discussed.

CHAPTER I

INTRODUCTION

Origin of the Problem

Teacher effectiveness has long been a concern of educational researchers (Morsh & Wilder, 1954; Getzels & Jackson, 1963; Medley & Mitzel, 1963; Ryans, 1963; Gage, 1965; Flanders, 1969); but with the movement of educational systems toward "accountability," teacher effectiveness has become a matter of first priority (Ryan, 1968; Lessinger & Allen, 1969; Sorensen, 1970; Silberman, 1969; Combs, 1972; Thatcher, 1972). The major question with regard to the relationship of teacher effectiveness and accountability is that of measurement--what it is that we should measure and how to measure it.

Flanders has suggested what it is that we should be measuring. He focuses on "relationships between the characteristics of teachers, teaching acts, and their effects on the educational outcomes of classroom teaching" (Flanders, 1969, p. 1423). The general purpose of the present study was to compare two approaches designed to measure the variables emphasized by Flanders: namely, the approach represented by attempts at systematic classroom observation and the approach using holistic measures related to characteristics of the persons in the process.

Background of the Problem

Attempts to establish a predictive relationship between student growth and teacher behavior have been undertaken by many researchers using procedures that have ranged in variety from Horn's (1914) attempt to account for the "distribution of opportunity for participation" by marking circles and squares on a seating chart, through Anderson's (1946) classification of teachers into "dominative vs. integrative" types, to the current preoccupation with systematic classroom observation instruments (Flanders, 1967). Of all these attempts, the most effective to date have been those studies using classroom observation systems (Flanders, 1969). Even these, however, have not been fully satisfactory.

Teacher effectiveness as defined by Flanders implies a straight line paradigm, as in Figure 1.

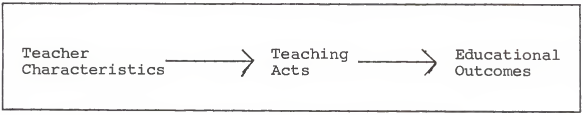


Figure 1. Teacher Effectiveness Paradigm

When researchers have dealt with *any two* of the three elements of the paradigm, the relationships have been upheld. However, when researchers have dealt with *all three* elements in the paradigm, the expected straight-line relationships have *not* been demonstrated (Sarason, 1962; Harvey, 1967; Soar, 1968; Smith & Kleine, 1969). In discussing his 1968 study, Soar commented:

> That is, that the presage-product relationships[teacher characteristics/educational outcomes] only relatively infrequently can be accounted for by means of the intervening, inter-related process [teaching acts]. The implication suggested is that the process measures [measures of teaching acts or instructional behaviors] used here have missed aspects of process to which the pupils in the classroom were responding. [Underlining by present writer]. (Soar, 1968, pp. 8-9)

Purpose of the Study

The specific purpose of the present study, then, was to determine if teacher effectiveness, based on student growth, could best be measured by a specific behavior or holistic approach to the problem. In the opinion of this researcher, the classroom observation systems approach deals solely with isolated pieces of behavior and, therefore, misses those process variables hypothesized by Soar. On the other hand, the Aspy scales are more holistic measures, and consequently, more likely to include such variables. It is possible that the missed aspects of process to which pupils in a classroom are responding are the perceptions and purposes of the persons engaged in the process. From this viewpoint, the paradigm shown in Figure 1 would have to be expanded to include two additional mediational elements, as in Figure 2.

As presented, this expanded paradigm proposes that the conversion of teacher characteristics into teaching acts is mediated (interpreted, influenced, and directed) by the teacher's purposes in teaching and by his perceptions of the

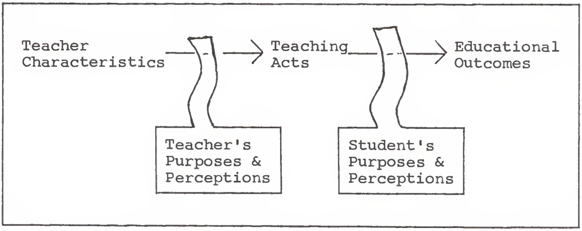

Teacher Characteristics
Teaching Acts
Educational Outcomes
Teacher's Purposes & Perceptions
Student's Purposes & Perceptions

Figure 2. Expanded Teacher Effectiveness Paradigm

pupils and of the instructional situation. Similarly, the effect of teaching acts on educational outcomes is mediated by the student's purposes and by his perceptions of the teacher and of teaching acts.

The hypothesis of the study has been that *holistic measures* of classroom process (which reflect the influence of the mediational elements of the expanded paradigm) would be more highly predictive of educational outcomes than would *specific-behavior measures* of classroom process (which do not take these mediational elements into account). That is, in order more accurately to predict educational outcomes, it is *not enough* to measure *only* the specific behaviors occurring within the teaching process; rather, it is imperative to take into account the student's perceptions of teaching acts and of the teacher's purposes.

Definition of Terms

The four major terms of the above hypothesis (educational outcomes, specific-behavior measures, holistic measures, more highly predictive) are defined below.

Educational Outcomes

For purposes of this study, educational outcomes were defined as the change from pretestings to posttestings across one school year (from September to May) of student scores on each of the following indices:

Index	Instrument
Achievement in-- Vocabulary Reading Spelling Arithmetic	Vocabulary, Reading, Spelling, and Arithmetic Concepts: sub-tests of the Iowa Tests of Basic Skills (Lindquist & Hieronymus, 1956)
Self-Concept factors-- Teacher-School Physical Appearance Interpersonal Adequacy Autonomy	Factor scores from the How I See Myself Scale (Gordon, 1968)
Motivation	Junior Index of Motivation (Frymier, 1965)
Dependence-Proneness	The Dependence-Proneness Scale (Flanders, Anderson & Amidon, 1961)
Anxiety as shown on-- Anxiety Scale Lie Scale	"A" and "L" scales of the Children's Manifest Anxiety Scale (Castanada, McCandless & Palermo, 1956)
Creativity Factors-- Fluidity Flexibility Originality	Three sub-tests of the Torrance Tests of Creative Thinking (Torrance, 1967)
Locus of Control	Social Reaction Inventory (Rotter, 1966)

Specific-Behavior Measures

Specific-behavior measures are measures of classroom process which occur at a very concrete level of abstraction; i.e., the level of deciding into which predetermined category a single behavior falls. Once this has been determined, a tally is made. Later, the records of this coding process are counted, totaled, or combined in some fashion by actuarial procedures to yield a *characteristic pattern of the recurrence of specific behaviors*. The only inference occurring is at the level of categorizing a single behavior.

The specific-behavior measures used in this study were two systematic classroom observation instruments: (1) the Reciprocal Category System (RCS) and (2) the Florida Climate and Control System (FLACCS). In terms of the expanded paradigm presented in Figure 2, neither of these two systematic classroom observation instruments (RCS or FLACCS) considers the mediational elements in the paradigm; i.e., they *do not* consider the individual assignment of meanings by students or teacher. The specific behaviors are coded into categories on the basis of the normative definition of the behavior. For example, saying "good" is categorized as *praise*, regardless of how the saying of it is received by the student.

Holistic Measures

Holistic measures are measures of classroom process which occur at a considerably higher level of inference than do specific-behavior measures. Instead of dealing with one single behavior, they take into account all the behaviors in the awareness of the observer at a specific instant of time and the totality of these behaviors is used to infer something about the nature of the perceptions held at that point of time by the individual(s) being observed. This process is repeated many times during the period of measurement. Depending on whether the totality of behaviors at many instants of time during that period yields the same or different inferences about the perceptions of the individual, then a further inference is made that this is a more or less permanent characteristic. The measurement instrument is consulted and a number is selected to represent the level (intensity) of that characteristic of the phenomenal field of the individual. This number is then recorded.

So in using holistic measures, all of the counting and combining of behaviors, as well as the inferences as to the meaning of the behaviors, takes place within the rater. The procedure clearly occurs at a higher level of inference than the procedure for specific-behavior measures.

The holistic measures used in this study were scores from the Aspy Rating Scales of Meaning, Genuineness, and Respect. In terms of the expanded paradigm presented in Figure 2, the Aspy Scales do consider the mediational elements in the paradigm; i.e., they do consider the individual assignment of meanings by students and teacher. According to Harris, when using "subjectivist" (holistic) measures, the variable is operationally defined by "designating the rater and the rating procedures" (1963, p. 20). A rater using the Aspy scales is trained to observe behavior in triads: (1) a student production, (2) the teacher's response to the student's product, and (3) the subsequent reaction by that student or by the next student who interacts with the teacher. As each behavior of the triad occurs, the rater makes an inference as to the purpose of the performer; and at each subsequent behavior in the triad, he makes an inference as to how the preceding behavior of the "other" person was received by the current performer. For example, a behavior triad might go like this:

a. A student (Robert) makes an incorrect response to a problem previously posed by the teacher.

b. The teacher responds, "This is a tough problem. Mary, would you try it?"

c. After a short pause, Mary responds--also inaccurately.

A rater using Aspy's Rating Scale for Respect would make the following discriminations at each phase of the above triad:

a. Did the problem call for a "fact" or a "thinking" response? Was the student's purpose appropriate to the demands of the problem; i.e., was he trying to answer at the level requested?

b. Was the teacher's purpose to communicate respect for the students' abilities to solve a hard problem or to communicate that she excuses their failure to solve the problem because it is a hard one and she did not really expect them to be able to do it? Did she receive Robert's answer as an expected failure or as a worthwhile endeavor to attempt a hard problem?

c. Did Mary receive the teacher's response as a communication of respect for her ability to solve the problem or as expectation of failure? Was her response appropriate to the demands of the problem and/or to the level of expectation communicated by the teacher?

Similarly, a rater using Aspy's Rating Scale for Meaning would make these discriminations at each phase of the triad:

a. How did the student feel about his answer? Was he purposefully trying to communicate the meanings of his experience?

b. Did the teacher's tone and choice of words communicate that she understands how the student felt about his answer? That is, (1) did she understand and (2) was her purpose to communicate to the student that she understood and accepted his feelings about his answer?

c. How did Mary feel about her answer? Did she seem to expect that her answer would be accepted by the teacher and her feelings about it understood and accepted? That is, was she purposefully trying to express the meanings of her experience?

And, finally, a rater using Aspy's Rating Scale for Genuineness would make these discriminations at each phase of the triad:

a. Did the student seem to expect his answer to be received in the context of who he is as a person, or in the context of his role as a student, or _as if_ divorced of relationship to the persons in the situation?

b. Does the teacher receive the answer within the context of the person or _as if_ it were divorced from the persons in the situation? Does she respond genuinely and spontaneously from herself as a person to another person,or as a person to a role, a role to a role, or a role to an object?

c. Does Mary seem to feel that she was called upon by the teacher as "Mary" or as "student" or as "answer-producing mechanism"?

Clearly, then, the rater characterizes (rates) not only the teacher's action, per se, but also makes inferences as to the teacher's purpose for the action and how that action was received by the student(s). This process is repeated for all triads during the time period rated, with the end behavior of one triad being the initial behavior of the next triad. Thus, the actual numerical rating which is finally assigned to the teacher to represent the level of Meaning, Respect, or Genuineness for that time period is an estimate by the rater of the general degree to which the teacher communicates (transmits and is received as having transmitted) more or less of the characteristic being assessed. In short, the Aspy Scales do take into account the effects of the mediational elements of the expanded paradigm.

More Highly Predictive

More highly predictive refers to the results of comparing the effectiveness of the two kinds of measures in accounting for the variability of student change scores on an educational outcome. This procedure occurred in two steps: (1) construction of two regression models (one model using holistic measures as predictors, and one model using specific-behavior measures as predictors) to predict each educational outcome and (2) comparison of the multiple R's obtained (amount of variance accounted for) by each model for each of the educational outcome indices.

Rationale for the Study

The study hypothesis of greater predictive power for holistic measures as opposed to specific-behavior measures of classroom process is based upon the perceptual psychology of Combs and Syngg (1959). In this framework, behavior is determined by the perceptual field of the behaver. Perceptions within that field are organized in more or less centrality to the self and those perceptions which are more central to the self are of more importance than perceptions in the periphery of the phenomenal field.

> . . . while the behavior of the individual is always the result of the total organization of the perceptual field, certain differentiations become so important in the production of behavior as to serve as effective guides to what the individual will or will not do in any circumstance. These differentiations are the phenomenal self, the individual's perceptions of the world about him, and the goals, values, and techniques he has differentiated in the course of his experiences. (Combs & Syngg, 1959, p. 447)

Therefore, in perceptual psychology, prediction of an individual's behavior is based upon understanding the meaning *to him* of the events in his phenomenal field. This means that if prediction from the teacher's acts to the students' behaviors (educational outcomes) is to be accurate, it must somehow take into account the meanings which the student assigns to the teacher's purposes and behaviors (teaching acts).

This led, therefore, to the study hypothesis: Because the set of Aspy Rating Scales is a holistic instrument which takes into account to some degree the student's perceptions of classroom process, it should be more highly predictive of subsequent student behavior (educational outcomes) than systematic classroom observation instruments such as RCS and FLACCS which are only a measure of specific behaviors within the classroom process.

CHAPTER II

REVIEW OF LITERATURE

Research on Teacher Effectiveness

Teacher effectiveness research prior to 1950 was not productive. After reviewing more than 1,000 studies published in the period between 1900-1952, Morsh and Wilder concluded:

> No single, specific, observable teacher act has yet been found whose frequency or percent of occurrence is invariably and significantly correlated with student achievement. (1954, p. 4)

Reasons assigned for the irrelevancy of the earlier research include (a) invalid criteria of teaching effectiveness, (b) lack of objective measures of teacher behavior, (c) research done in a theoretical vacuum, (d) lack of adequate research design, and (e) problems of definition (Gage, 1965; Getzels & Jackson, 1963; Medley & Mitzel, 1963).

In the last two decades, however, some trends of greater promise in measuring teacher effectiveness have become evident. They include (1) improvement in instrumentation and observational systems, (2) greater attention to theory and problems of definition, (3) a shift from subjective evaluations to more objective measures, and (4) a shift from attention to teacher effectiveness, as such, to increased attention to the fundamental problem of the description of teacher behavior and to the conditions with which various teacher behaviors are correlated with each other and with criterion variables (Ryans, 1963; Flanders, 1969).

Research, then, under these "improved" conditions has been conducted along two divergent lines--one line concerned primarily with holistic variables or "who" the effective teacher is and the other line emphasizing specific behavior variables or what it is that the teacher does which is growth-producing in students. Both lines of research have yielded dividends.

Holistic Variables

On the one hand, those investigators concerned with holistic variables have demonstrated that "who" the teacher is, is related both to effectiveness as a process measure and to student outcomes. For example, significant relationships have been demonstrated between ratings of teacher effectiveness and the teachers' MMPI scores (Moore & Cole, 1957), factors in Rorschach protocols (Cooper & Lewis, 1951), MTAI scores (Cook et al., 1956; Leeds, 1956), adequacy of psychological functioning (Guba & Getzels, 1955), and profiles on Cattell's 16PF (Lamke, 1951). Willard (1955) found a significant relationship between the values of teachers and the learning experiences provided in their classrooms. Still other investigators found significant relations between teacher characteristics and such student products as achievement gains, liking for school, decreased anxiety, increased divergent thinking, cooperation, involvement, activity level, helpfulness, decreased nurturance seeking, and reading gains of retarded readers (Wallen, 1966; Heil, 1962; Heil and Washburne, 1962; Harvey, 1967; Lipton, 1967).

Of particular note within this line of investigation is the work instigated at the University of Florida which demonstrates the great dividends to be attained from attention to theory and problems of definition--the second trend "of great promise" mentioned by Ryans (1963). Subsequent to a seminar which explored the idea that effective helpers had some perceptions in common, several researchers investigated the perceptual dimensions defined by the faculty and graduate students who had participated in the seminar (Combs et al., 1969).

Several studies were particularly concerned with teaching. Gooding (1964), Usher (1966), and Vonk (1970) demonstrated that effective teachers were characterized by similar perceptual organizations; their studies delineated aspects of that organization and its holistic nature. Courson (1965) tested the reliability of inference as a research tool and demonstrated both between observers and within observers reliability at significance levels ranging from .025 to .005. Brown (1970), Choy (1969), and

and Doyle (1969) identified specific aspects of the teacher's perceptual organization which were related to effective teaching. In addition, Choy (1969) found that teachers' conceptual systems orientation was related to their perceptual orientation and to teaching effectiveness; Brown (1970) found that elementary and secondary teachers differed in certain aspects of their perceptual organizations; Doyle (1969), Choy (1969), and Morgenstern (1969) demonstrated that agreement on teacher effectiveness using global inferential ratings can be obtained among administrators, colleagues, and students, while Morgenstern also found that global subjective ratings of teacher effectiveness were correlated with level and direction of verbal self-disclosure and tactual body-accessibility. Aspy (1968) attested that the teacher's self-concept was correlated with the achievement of students.

Most recently, this line of research has culminated in the evaluation of the New Elementary Program (NEP) at the University of Florida. This program is an experiment in systematic theory-based curriculum innovation. That is, the program is an essentially individualized approach to the preparation of elementary teachers with both the curriculum and the professor/student relationships based on theories, principles, and research from perceptual psychology. In evaluation of the program, Wass et al. (1974) found significant ($p<.01$) differences between NEP and control groups on

"Perceptions of Others as Able-Unable," on "Perceptions of Goals as Larger or Smaller," and on the total scores of the Perceptual Dimensions Scale. All other differences were also in favor of the NEP group, although only approaching significance.

Specific-Behavior Variables

Investigators emphasizing specific-behavior variables (generally measured through the use of systematic classroom observation instruments) have established several interesting, replicable, and fairly stable correlations. Some examples include:

(a) "Responsive" Teacher Behavior (giving praise and using pupil ideas) is related positively to achievement in language usage, social study skills, arithmetic computation, problem solving, suspension of judgment, and science knowledge. (Morrison, 1966; LaShier & Westmeyer, 1967)

(b) Teachers talk between 55 and 75 percent of the time in such a fashion that pupil talk is primarily in response to the initiative of the teacher and deals with verbal factual recall. Less than 1 percent of the classroom interactions are initiated by pupils. (Flanders, 1969)

(c) Teachers are more alike than they are different and the less effective teachers are more alike and less flexible and probably easier to identify than more effective teachers are. (Flanders, 1965)

(d) Significant negative relationships were obtained between some form of teacher criticism of students and measures of student achievement in five studies. (Anthony, 1967; Flanders, 1970; Harris, 1968; Soar, 1966a; Wallen, 1966) Two studies indicated that the stronger the form of criticism, the larger the negative correlation with achievement. (Hunter, 1969; Perkins, 1965)

(e) "Responsive" Teacher Behavior is related to teacher self-concept, security, values, attitudes toward students, and emotional adjustment. (Davies, 1961; Soar, 1966b)

Summary

To summarize, many investigators from a variety of viewpoints have established that measures of the personal characteristics of teachers are related to process measures, i.e., measures of the events occurring during the teaching act. Other investigators have demonstrated that teacher characteristics are related to educational outcomes. Still other researchers have shown evidence that process measures are predictive of educational outcome measures. This would seem to yield the paradigm presented in Figure 1 of Chapter I. Implied in that paradigm is a straight-line relationship of influence or association passing from teacher characteristics through teaching acts to educational outcome variables.

Researchers have been able to demonstrate relationships between each pair of any two of the three elements of the paradigm. That is, taking the elements of the paradigm only two at a time, researchers have been able to demonstrate relationships between,

1. teacher characteristics and teaching acts (Davies, 1961; Campbell, 1971; Gooding, 1964; Soar, 1966a; Usher, 1966; Wass et al., 1974),

2. teacher characteristics and educational outcomes (Soar, 1966a; Wallen, 1966; Smith & Kleine, 1969), and

3. teaching acts and educational outcomes (Davies, 1961; Flanders, 1965, 1969; LaShier & Westmeyer, 1967; Perkins, 1965; Morrison, 1966; Soar, 1966b; Anthony, 1967; Harris, 1968; Hunter, 1969).

However, the straight-line relationship passing through all three elements of the paradigm has not been substantiated in the few studies (Soar, 1968; Smith & Klein, 1969; Sarason, 1962; Harvey, 1967) that have tried to investigate the relationships of all three elements in the paradigm. In summarizing their study, Smith & Klein commented, " . . . it is apparent that the proposed chain which led from teacher complexity through [communication of] teacher awareness to pupil esteem was not supported by our data" (1969, p. 251). And in regard to the results of his study, Soar commented,

> A second implication of these results may be that numbers of presage characteristics of the teacher . . . do relate to product changes [educational outcomes] in ways which the process [teaching acts] measures do not reflect . . . . There seems little question but that prediction of pupil growth from process measures is more successful than from presage measures, but at the same time it is clear that presage characteristics that relate to pupil change are not consistently reflected in process measures. (1968, pp. 8-9)

Soar went on to recommend,

> Lacking important data on all fronts, as we now appear to do, the more fruitful strategy would appear to be that of emphasizing process-product

> relations; and using these, in turn, to alter training programs, and hire teachers, and finally, to develop process-related presage measures. (1968, p. 10)

One of the "fronts" on which important data is lacking is in the area of comparisons between the variables being used in the two distinct lines of teacher effectiveness research discussed above. The present study attempted to provide some of that "lacking" data by making a direct comparison of the effectiveness of process variables from each of the two research approaches. That is, it compared holistic measures of process variables related to characteristics of the persons in the process with specific-behavior measures of the process itself.

Background of the Selected Holistic Instrument--the Aspy Scales for Rating Interpersonal Processes

The Aspy scales of Meaning, Genuineness, and Respect are based on the theoretical positions of Combs (Combs & Snygg, 1959; Combs et al., 1969) and Rogers (1961) and on the empirical investigations of Carkhuff (1970).

Theoretical Base

Research studies (Combs & Soper, 1963; Benton, 1964; Gooding, 1964; Usher, 1966; Dickman, 1967; Vonk, 1970), which have investigated the perceptual organization of effective helpers, support the idea that effective helpers, including teachers, have a number of similar perceptual characteristics. They also established the concept that a teacher having these characteristics would differentiate techniques* in such a way that he would act as an "intelligent human being using himself, his knowledge, and the resources at hand to solve the problems for which he is responsible" (Combs, 1965, p. 8)

Rogers (1961) postulated that one of the ways (techniques) in which effective helpers in counseling used themselves was by establishing a facilitative relationship with the client. Such a relationship would be characterized by the presence of empathy, congruence, and positive regard offered by the helper to the client. Rogers also postulated that teacher-provided levels of these same three conditions were directly related to learning outcomes:

> . . . To the extent that the teacher creates such a relationship with his class, the student will become a self-initiated learner, more original, more self-disciplined, less anxious and other-directed. (1961, p. 37)

Early Studies

Using Rogers' theoretical formulations, Truax and Carkhuff (1967) and Carkhuff and Berenson (1967) devised and validated a set of scales and procedures that could be used to measure the levels of facilitative interpersonal conditions. Although these scales for measuring empathy, congruence, and positive regard were not developed specifically for assessing classroom interaction, educational researchers have been able to use them to establish that the levels of interpersonal conditions offered by the teacher were significantly and positively related to such pupil outcomes as the following: cognitive growth measured by standardized tests (Aspy, 1967; Aspy, 1969; Aspy & Hadlock, 1967), student performance (Truax & Tatum, 1966), decreased truancy rates (Aspy & Hadlock, 1967), and grade-point average (Pierce, 1966).

---

*As used here, "technique" is defined as a method of reaching goals and satisfying needs by moving through a characteristic chain of sub-goals which have been differentiated from the perceptual field of the individual. The specific nature of the sub-goals in the chain are not fixed, but the overall meaning of the technique is stable and can be identified in terms of its function in the maintenance or enhancement of the self. Once differentiated, "techniques become an integral part of the organization of the individual's perceptual field" and "often have a high degree of stability and provide a measure of predictability for behavior." (Combs and Snygg, 1959, pp. 114-120)

Most of the investigations above used the Carkhuff scales as a measure of some teacher characteristic which was being expressed in the levels of interpersonal conditions offered to the students. In 1971, Dellow investigated this assumption by comparing the variables used in the perceptual approach for inferring the presage characteristics of the teacher (i.e., his phenomenal field) to the variables used in the facilitating conditions approach. On the basis of "positive but low correlations between perceptual and behavioral variables" and "failure to find a significant relationship [canonical correlation] between perceptual variables as a set and the behavioral variables as a set," (p, 80) he concluded that the two approaches may measure different facets of a teacher's behavior. Additionally, he concluded that the scales did not apply directly from the counseling situation to the classroom situation so that the raters had to establish a "standard" interpretation for the classroom activities.

Scales Reformulated

This set of scales has now been reformulated (Aspy, 1972; Aspy & Roebuck, 1973, 1975; Aspy, Black & Roebuck, 1972) specifically for tapping the levels of interpersonal conditions as they are manifested in classroom situations. In the reformulation, the scale for measuring levels of empathy was renamed Meaning; congruence became Genuineness; and positive regard became Respect. (See Appendix I for text of the Scales.)

The changes seemed necessary (1) to make the language of the scales more appropriate to the classroom context and (2) to relate the scales more directly to the theoretical position of perceptualistic psychology in which teaching is seen as a human interrelationship for the purpose of helping people discover the personal meaning of information (Combs & Snygg, 1959). In the reformulation of the scales, rater-training procedures were standardized (see discussion of observation in triads, Chapter I, p. 8) so that the scales are clearly holistic measures which include inference as to the perceptions of the individuals in the process as well as consideration of the actual behavior being emitted.

The word in the scales which keys to this process of attention by the rater to the individual assignment of meaning is "communicate" as communication refers both to the transmission and to the reception of the message. Thus, the rater observes not only what the teacher transmits, but what the student receives, i.e., the meaning assigned by the student to the teacher's behavior.

The scales in their new formulation have been used in a series of studies and have yielded replicable relationships with dependent variables. (See discussion of Instruments in Chapter III.)

Background of Selected Specific-Behavior Instruments--RCS and FLACCS Systematic Classroom Observation Systems

The efforts to quantify teacher and student behaviors through systematic classroom observation had its beginning with Horn's (1914) suggestion that classroom visitors could record "recitation or requests for recitation" by placing a small circle or square in the appropriate space on a seating chart. Puckett (1928) elaborated this codification system by adding dots, lines, and other symbols to the circles and squares proposed by Horn. These early attempts at systematizing the observation of classroom processes were extended into the area of interaction analysis (or, as it was then called, "classroom climate") by Thomas (1929) in studies of nursery-school behavior.

Early Studies

Of the early studies in interaction analysis, the series carried out by Anderson in the early 1940's has become a classic. These studies established a mode of data analysis which is still the most widespread methodology for using the information obtained by interaction analysis. In this mode, the data are used to classify teachers into one or more "types" along some dimension of the "climate" being measured. For Anderson (1939, 1946), this dimension was "dominative vs. integrative." Other researchers have used other dimensions, such as "authoritarian vs. democratic vs. laissez-faire" (Lippet & White, 1943), "teacher-centered vs. student-centered" (Withall, 1951; Perkins, 1951), "preclusive vs. inclusive" (Cogan, 1958). These types are then used as discrete groups for the further analysis of outcome data. For example, the amount of self-initiated work done by pupils of "inclusive" teachers is compared with that done by students of "preclusive" teachers in order to determine the relative effectiveness of the two types of teachers.

This method of data analysis has two drawbacks. First, it leads to a loss of scientific objectivity because of the use of value-packed verbal labels for the teacher types. According to Flanders, the precise meaning of "climate" when commonly used is

> . . . seldom clear, just as its synonyms "morale," "rapport," and "emotional tone" are also ambiguous. To have any meaning at all, the word must always be qualified by an adjective, and it is in the choice

> of adjectives that researchers reduce their scientific integrity by losing their objectivity . . . . (1967, p. 104)

The second drawback to this system is that it can only be used to provide information about the general effects of teacher behavior over long periods of time. In so doing, it ignores the flexibility of teaching behavior in short-term patterns and acts under the assumption that the teacher is a static entity rather than a dynamic system in process.

Later Development

Following the 1939 and 1946 studies by Anderson, Withall (1951), Medley & Mitzel (1958), and Hughes (1959) developed other classroom observation systems which were simple to use and, with the accompanying development of point-time sampling techniques, provided good reliability among observers. However, the data yielded and the analysis made of those data were still little more sophisticated than those of Anderson's studies.

The next major advance came with the publication of Flanders' (1965) system of verbal interaction analysis. With Flanders' system not only the quantity but the sequence of classroom interactions could be codified. The utility of this codification system has been demonstrated by the significant and replicable results it has yielded. Although not without its drawbacks, Flanders' system comes very close to meeting the need for "a dynamic explanation of how short-term patterns of teacher influence affect momentary situations so that the flexibility of the teacher's behavior is taken into account" (Flanders, 1967, p. 107). It provides the codification system that preserves the data (both quantitative and sequential) that is necessary for such analysis of short-term patterns.

The specific classroom observation measures used in this study were the Reciprocal Category System (RCS) and the Florida Climate and Control System (FLACCS). A modification of Flanders' Interaction Analysis, the RCS adds categories to enable the researcher (1) to distinguish between teacher lectures that are responsive to pupils and those that are teacher-initiated; (2) to distinguish between criticism and corrective feed-back; and (3) to identify pupil questions. The resulting data can be converted to Flanders' system by pooling categories, if desired, so that it retains comparability with previous research.

FLACCS is a further refinement of the SCOR (Soar, 1966a) which was assembled from several revisions of the Medley & Mitzel (1958) observation schedule and record. Items were selected primarily to represent aspects of classroom process not identified by Flanders' Interaction Analysis. In addition, the Hostility-Affection Schedule of Fowler (1962) was included for its emphasis on expression of affect: teacher/pupil, verbal/non-verbal, supportive/non-supportive.

Both the RCS and the FLACCS are in frequent usage by researchers (see discussion of Instruments in Chapter III) and have proved their efficacy in predicting student outcomes. The selection of these particular specific-behavior instruments for comparison with the holistic Aspy Scales provided a test of the relative power of the two approaches in predicting student outcomes which was as fair as it was possible to devise. That is, matching the predictive power of one holistic instrument against the combined predictive power of two specific-behavior instruments biases the study in favor of the specific-behavior measures. But this procedure seemed dictated by the fact that no one specific-behavior instrument tapped all of the dimensions considered by the holistic measures. The two selected instruments (RCS and FLACCS) between them, however, did count behaviors in all of the dimensions tapped by the holistic measures, i.e., instructional interactions, expression of positive and negative affect, levels of cognitive functioning, and interpersonal relationships. (For additional discussion of this point, see "Bias of the Study," in Chapter III.)

Research Related to Study Assumptions

In conducting this research, several assumptions have been made. Each of these assumptions is specified below and research in support of the validity of the assumption is reviewed.

Assumption 1

The basic assumption of the study is that the student's perceptions of the teacher's behavior will have a directing and governing influence on his subsequent behavior. This principle is well supported in perceptualistic psychology (Combs & Snygg, 1959). Additionally, Pendergrass (1971) demonstrated that students' perceptions of self-adequacy were related to their perceptions of threat. Malpass (1953) found that student perceptions of school (scores on a sentence completion test, a ten-card apperception test, and a composition on their feelings about school) were significantly related to their end-of-semester grades and less highly related to arithmetic achievement tests.

Although operating from a different theoretical framework, Levy (1961) conducted a series of studies that demonstrated the potency of the perceiver in assigning meanings to what he sees and thus influencing his subsequent behavior, with the main mechanism being "interpretation" rather than "selection" of cues. More directly supportive to the area of the current research was the study of Schmuck (1966) in which he found that the behavior of pupils who perceived their teachers as understanding them was characterized by "a more diffuse liking structure" among themselves (Chi Square = 4.5; $p < .05$) and the expression of "more positive attitudes toward school" than was the behavior of pupils who perceived their teachers as not understanding them.

Assumption 2

A necessary assumption of this study is that there is at least some commonality among students within a single classroom as to their perceptions of classroom process.[*] This idea was also supported by Schmuck's 1966 study in that he found significant differences between classroom groups of students in their perceptions of their teachers. Meyers (1971) also found significant differences to exist between perceptions of teachers reported by students in "open area" classrooms and those by students in "self-contained" classrooms.

When considering the total group of teachers, Gage et al. (1955) found no significant relationships between pupils' ratings of their teachers and the teachers' scores on measures of accuracy in identifying student (1) personal problems, (2) social relations, and (3) level of cognitive functioning. They did find significant differences between pupil ratings when the teachers were stratified by grade level, age, sex, or score on the Minnesota Teacher Attitude Inventory. Gillett (1965) demonstrated that changes in student teachers' acceptance of pupils were significantly related ($p < .05$) to pupil self-reports of their acceptance of the student teacher (scores on a Teacher Rating Scale).

---

[*] Note: This assumption is derived from a principle of perceptual psychology: "The presence of common meanings in the perceptual fields of different persons makes communication possible. It also makes possible the understanding and prediction of the behavior of other people." (Combs & Snygg, 1959, p. 35)

Perhaps the most definitive study in this area is that of Wright and Sherman (1965). They analyzed the student's ratings of their teachers in two ways:

> To find out the level of general agreement among their perceptions of the teacher, the average rating of the teacher for each class of children was calculated for each item. Then, to extract the common themes underlying children's disagreement in their perceptions of the teacher, the principal components of the covariance among pupils' ratings were obtained for each class. (p. 92)

In general, Wright and Sherman found that students tend to agree on their ratings of role-specific qualities of the teacher while at the same time expressing individual differences in their rating of qualities relevant to the teacher's personal relations with the students.

Assumption 3

The assumption that students are capable of making reliable judgments as to the interpersonal aspects of classroom process and that they do make such judgments is supported by several studies. Emmerling (1961) reported that teachers who had been classified as being more "open" were perceived by their students as significantly ($p < .001$ to .0001) more empathically understanding, more congruent, and more pupil-centered than teachers not so classified. Bledsoe, Brown, and Strickland (1971) reviewed the research on student ratings of teachers and reported that they were at least as reliable as the "better educational tests presently available." (p. 119) Dixon and Morse (1961) asserted,

> Our results indicate that pupils do have different perceptions of the individual's ability to develop good interpersonal relationships . . . as far as the pupils were concerned, the student teachers who had "good" empathy were also better teachers. (pp. 328-329)

And Silberman concluded that

> . . . teachers' attitudes are generally revealed in their actions, in spite of many forces operating to contain their expressions . . . the students who receive them are aware of most behavioral expressions of their teachers' attitudes . . . many such behaviors aimed at individual students are visible to other students in the class, as well. . . . Students were able to predict better than chance the relative amount of contact, negative evaluation, and acquiescence they received . . . and were also able to predict how much negative evaluation and acquiescence their classmates received. (1969, pp. 405-407)

Assumption 4

The last assumption of this study is that there is not a one-to-one correspondence between students' attribution of meaning to behavioral events and those events as counted or recorded by systematic classroom observation measures. This assumption was supported by three studies. Hays (1970) found that student perceptions did not correspond closely with communication climate classifications by classroom observers. Tuckman (1970, p. 39) hypothesized that the reason why the student judgments in his study did not conform more closely to observer's ratings might be a function of the fact that

> (1) students based their judgments on a different set of teacher behaviors than did the observers; i.e., the same behavior observed by different judges was

> rated differently or (2) students based their judgments on their overall experiences within the classroom; i.e., a host of different behaviors had the same connotative [inferential] meaning and were therefore similarly rated. (1970, p. 39)

Gray, Graubard, and Rosenberg stated that

> While the students quickly learned to score negative behavior, they were seldom able to recognize positive behavior in their teacher-clients. Without the knowledge of the teachers or of the students, trained adult aides also kept records of teacher behavior in the classes . . . . the aides recorded substantially more instances of positive teacher behavior than did the students. For example, an aide reported that a teacher had praised a child, but the child reported that the teacher had chewed him out. . . . determined through closer monitoring that the aides were more accurate [in terms of the verbal content of the teacher's communication]. . . . the students were unable to recognize positive teacher behavior because they were accustomed to failure and negative treatment. (1974, p. 44)

And French (1970, p. 194) summed it all up in his statement: "We have a communication problem when we assume that meanings are in words rather than in people."

CHAPTER III

DESIGN OF THE STUDY

Overall Design of the Study

The present study was essentially a single-group "comparison correlational survey" (Fox, 1969, p. 46) testing hypothesized relationships among the data collected. That is, the intent of this study was to conduct comparisons among multiple sets of data collected from a single sample in order to make a judgment of "difference" or of "no difference" between certain sets of the collected data. In this case, the judgment which would be made was whether there was a difference in the degree to which two sets of data (one set representing the holistic approach to the measurement of teaching acts and the other set representing the specific-behavior approach to the measurement of teaching acts) were correlated with sixteen other sets of data representing changes from pretestings to posttestings on specified student outcome indices. A split-sample technique was used to create independence for statistical testing of null hypotheses.

Hypothesis

The hypothesis of this study was that: Holistic measures (scores from Aspy Rating Scales) of classroom process will be more highly predictive of the variability in change on certain specified Educational Outcome Indices[*] than will specific-behavior measures (factor scores from RCS and FLACCS Classroom Observational Systems).

Statement of Null Hypotheses

Translation of the above study hypothesis into form for testing in the statistical procedures yields sixteen null hypotheses, one for each outcome variable. They take the following form:

When a regression model using quadratic terms from Aspy Rating Scales as predictors is compared against a regression model using incomplete factor scores from RCS and FLACCS Classroom Observation Systems as predictors, the Aspy Rating Scales model will not account for a larger amount of variance in change than will the RCS and FLACCS model when the educational outcome index consists of student gain on

1. Reading Vocabulary Achievement
2. Reading Comprehension Achievement
3. Spelling Achievement
4. Arithmetic Achievement
5. Children's Manifest Anxiety, Scale A
6. Children's Manifest Anxiety, Scale L
7. Dependence-Proneness
8. Social Reaction Inventory
9. Junior Index of Motivation
10. How I See Myself, Factor: Teacher-School
11. How I See Myself, Factor: Physical Appearance
12. How I See Myself, Factor: Interpersonal Adequacy
13. How I See Myself, Factor: Autonomy
14. Creativity Factor: Fluidity
15. Creativity Factor: Flexibility
16. Creativity Factor: Originality

---

[*] Indices listed on p. 5

Sample

The study used data obtained from fifty fifth-grade classrooms in six counties near Gainesville, Florida. The samples included both Negro and Caucasian teachers and students from rural, suburban, and urban populations. Distribution of classrooms by county and school is presented in Table 1. Race and sex distribution of students by classroom is set forth in Table 2. Demographic data are presented in Table 3.

Table 1

DISTRIBUTION OF CLASSROOMS IN STUDY

County	School	Number of Classrooms
Alachua	Kirby-Smith	2
	Stephen Foster	3
Columbia	Five Points	2
	Fort White	1
	Niblack	1
Dixie	Cross City	4
	Oldtown	1
Duval	Arlington Heights	4
	Axoon	2
	Daniels	1
Gilchrist	Bell	1
	Trenton	2
Marion	Anthony	2
	Belleview	3
	Citra	1
	Dunnellon	2
	Fessenden	1
	Ft. McCoy	1
	Howard	5
	N. H. Jones	2
	North Marion	2
	Oakcrest	4
	Wyomina Park	3

Table 2

STUDENT N WITH RACE/SEX DISTRIBUTION BY CLASSROOM

County	Classroom No.	Total	Male White	Female White	Male Black	Female Black
Alachua	1	17	3	5	4	5
	2	18	5	4	7	2
	3	16	1	0	6	9
	4	12	0	0	3	9
	5	23	0	0	13	10
Columbia	1	22	2	0	9	11
	2	19	2	0	10	7
	3	18	8	4	1	5
	4	13	8	2	2	1
Dixie	1	13	2	2	3	6
	2	12	1	3	6	2
	3	10	0	1	3	6
	4	13	1	3	5	4
	5	14	0	0	8	6
Duval	1	12	0	1	7	4
	2	17	0	0	10	7
	3	21	1	0	10	10
	4	20	0	0	10	10
	5	23	3	3	6	11
	6	13	1	2	7	3
	7	13	5	7	1	0
Gilchrist	1	12	2	0	4	6
	2	10	0	2	3	5
	3	15	1	2	5	7
Marion	1	13	2	2	5	4
	2	13	0	2	7	4
	3	16	2	4	3	7
	4	20	4	2	4	10
	5	20	2	4	6	8
	6	10	1	4	3	2
	7	20	3	2	7	8
	8	13	2	2	6	3
	9	13	1	10	2	0
	10	15	3	1	7	4
	11	14	2	7	0	5
	12	13	5	8	0	0
	13	15	8	7	0	0
	14	13	5	8	0	0
	15	13	3	10	0	0
	16	21	11	10	0	0
	17	24	9	15	0	0
	18	23	5	15	3	0

Table 2--continued

County	Classroom No.	Total	Male White	Female White	Male Black	Female Black
	19	19	4	8	4	3
	20	17	0	0	7	10
	21	23	0	0	11	12
	22	14	0	0	10	4
	23	14	1	0	7	6
	24	16	0	0	6	10
	25	13	0	0	9	4
Total		781	119	162	250	250

Table 3

STUDENT DEMOGRAPHIC DATA

Classification	Percent of Student Group
Disadvantaged Students (HEW Criteria)	34.8
Students Attending School in Community with Population of	
Less than 2,500	44.9
2,500 - 10,000	0.0
10,000 - 25,000	28.5
25,000 - 50,000	0.0
50,000 - 100,000	9.7
More than 100,000	16.7
Students Attending Schools with White/Black Ratios of	
100% White	14.5
90 - 99% White/ 1 - 10% Black	20.5
60 - 89% White/ 11 - 40% Black	35.1
40 - 59% White/ 41 - 60% Black	4.1
10 - 39% White/ 61 - 90% Black	8.6
1 - 9% White/ 91 - 99% Black	1.5
100% Black	15.3

Data Collection

Except for the scores for Aspy's Scales of Meaning, Genuineness, and Respect, the data in this study were collected under National Institute of Mental Health Research Grant No. 15891-01 to the Institute for Development of Human Resources, University of Florida, Gainesville, Florida. The data were made available to the current researcher by Robert S. Soar, principal investigator under the grant.

All data were collected within one school year. In September, pretest measures for each of the 16 educational outcome indices were taken on the pupils in the 50 classrooms sampled. At this time, an I.Q. test was also administered as I.Q. scores would later be used to adjust gain for individual differences.

Once a month from November to April, FLACCS systematic classroom observation procedures were carried out in all of the classrooms. At the same time that each observation was conducted, a tape recording was made of the instructional interaction. These recordings were subsequently used by raters at the Institute for Development of Human Resources to generate data from the RCS instrument. In May, posttest measures were taken from the pupils. The tape recordings used to generate data from the RCS instrument were preserved and made available to the current researcher for use in application of the Aspy Rating Scales. Further details of data collection for each instrument are discussed separately below.

Educational Outcome Indices

The pretests and posttests for each of the 16 measures of educational outcomes (listed on p. 35) were administered by the staff of the Institute for Development of Human Resources. The pretests were administered within a two-week period in September and the posttests in the first two weeks of May. All administrations took place in the students' usual classroom settings. The tests were scored by the test publishers and the data (individual student scores) returned on computer tapes to the Institute.

I.Q. Test

The nonverbal form of the Lorge-Thorndike Intelligence Test was administered in September. Procedures for administration and scoring were the same as those for the educational outcome indices.

Florida Climate and Control System (FLACCS)

Data with the FLACCS instrument were collected through a process of live systematic classroom observation. Each classroom was visited once a month from November through April by a team of observers who had been trained during September and October in a scheduled credit-bearing course at the University of Florida. Each team consisted of two observers who were graduate students, research assistants, or permanent members of the staff of the Institute for the Development of Human Resources. There were seven such teams and, within the constraints of scheduling, each classroom was visited by each team only once.

Each observation period consisted of a half-hour visit to the classroom. During the half-hour, the observers alternated five-minute periods of just observing, with five-minute periods of marking the instruments. Thus, the data reported from a single visit to a classroom consisted of the tallies for a total of fifteen minutes from a half-hour sample of instructional activities. Visits to the classroom were scheduled so that in the 6 observations, activities in 3 different subject-matter areas were sampled.

The marked instruments were returned to the Institute where they were prepared for computer accession. The data were submitted to factor analysis, using principal components factor extraction followed by varimax rotation of a series of numbers of factors. The number of factors rotated which seemed to offer the clearest set of interpretations was retained. Then, incomplete (non-orthogonal) factor scores were computed for each teacher by pooling the teacher's individual T-scores for the items which loaded ± .40 or above on each of the factors. These individual-teacher incomplete factor scores were the FLACCS data which were provided to the current researcher and used in the present study. The FLACCS factor structure with item loadings for each factor is provided in Appendix II.

Reciprocal Category System (RCS)

The observer teams carried small battery-powered audiotape recording equipment with them and, simultaneously with each observation, they made a recording of the instructional activity. The tape consisted of the recording of the three five-minute segments during which marking of the FLACCS instrument was occurring. The quality of the recording playback was up-graded by the use of a Cardiod microphone, Mallory batteries and chargers, "high-output" tape sensitive to faint signals, and stereo headsets rewired to function monaurally; consequently, the tapes could be coded for anything that a live observer would have been able to hear and understand in the classroom.

The tapes obtained were returned to the Institute where they were coded on the RCS by raters who had been trained during the early fall in a scheduled credited course at the University of Florida. The resulting RCS code sheets were subsequently made available for computer accession through keypunching. The RCS data were submitted to the same factor analysis procedures as those used with the FLACCS data. The individual-teacher incomplete RCS factor scores were the data which were provided to the current researcher and used in the present study. The RCS factor structure, with item loading for each factor, is furnished in Appendix II.

Aspy Scales of Meaning, Genuineness, and Respect

The tapes which had been used to generate the RCS data were made available to the current researcher. Three tapes for each teacher were selected at random from among the 6 observations for each teacher, with the only constraint being that each of the selected tapes should sample a different subject-matter area. When a drawn sample of three tapes was found to duplicate a subject-matter area, all three samples were returned to the pool and a new sampling was made.

From each of the three tapes selected for each teacher, two three-minute segments--one from near the beginning of the tape and one from near the end of the tape--were selected at random and then rated independently by three trained raters using Aspy's Scales of Meaning, Genuineness, and Respect. The segments from each tape were dubbed onto Master Tapes to ensure that the independent raters would rate identical segments for each teacher.

The raters were a professor and two graduate students in education. Rater-training was conducted by the current researcher and consisted of 15 hours of didactic and experiential training. The training ceased when 85% acceptable agreement (agreement within ±.3 of a scale level) was reached by all three raters.

Once acceptable levels of reliability had been reached, each rater independently rated all segments on the Master Tapes. The ratings for each scale by each rater on each segment were keypunched for computer accession at the Academic Computer Center of Northeast Louisiana University. In order to arrive at a scale score for each tape, the ratings (one for each of two segments by each of three raters) were averaged and the resultant mean became the teacher's score for that scale for that tape. Thus, each teacher had three scores (one for each scale) on each of three tapes, or a total of nine scores. These nine scores were the data used in the analyses reported in this study.

Instrumentation

The instruments used in the study included the Reciprocal Category System, the Florida Climate and Control System, the Aspy Rating Scales, an I.Q. test, and seven paper and pencil instruments administered in pretestings and posttestings to students. The sub-tests and the factors of the seven instruments (listed on p. 5) yielded a total of 16 educational outcome indices. All of the instruments selected for this study are in current usage by educational researchers. They are discussed separately below.

The Reciprocal Category System (RCS)

The Reciprocal Category System (Ober, Wood, & Roberts, 1968) is a modification of Flanders' Interaction Analysis Categories (Flanders, 1965). It offers about twice as much data as the original instrument with only a slight increase in the difficulty of the coder's tasks. In this modification, the seven teacher categories of Flanders' original instrument have been increased to nine, the two pupil categories have also been increased to nine, while Category 10 (Silence and Confusion) has been subdivided into two categories: Silence (10) and Confusion (20). The RCS instrument is shown in Appendix I.

Research with Flanders' Interaction Analysis Categories has consistently revealed significant relationships with changes on pupil outcomes (Flanders, 1965, 1970; Amidon & Flanders, 1961; Amidon & Hough, 1967; LaShier, 1967; Furst, 1967). As a result, it is the instrument most often used as a basis for expanding or synthesizing a system (Rosenshine & Furst, 1973). The RCS modification of Flanders' has also proved its usefulness in predicting student behavior (Ober, Wood, & Roberts, 1968; Soar, 1966a, 1973). Additionally, the data obtained with the RCS can be converted to Flanders' system by pooling categories so that it retains comparability with all of the research using the original instrument.

In the 1973 study, Soar reported inter-rater reliability figures of .84, .66, .55, .76, .97, .75, .82, .80, and .80 for the nine incomplete factor scores yielded by varimax rotation of the RCS data. Since the raters (coders), rater-training, and coding procedures in that study (which involved 289 K-2 classrooms) were essentially the same as for the RCS data reported in the current study, no further reliability studies were made.

Florida Climate and Control System (FLACCS)

The Florida Climate and Control System (Soar, Soar, & Ragosta, 1971) was assembled from several earlier instruments which had proved their utility in measuring specific aspects of classroom processes. The primary contributor to the modified instrument was the Florida Affective Categories System (Soar, 1971) which itself was a modification of the South Carolina Observation Record (Soar, 1966b). Items were also drawn from the Observation Schedule and Record (Medley & Mitzel, 1958); and the Hostility-Affection Schedule of Fowler (1962) was included almost in its entirety.

The resulting FLACCS instrument has two parts. The first part, with 84 items, is concerned with the teacher's control of the classroom and with pupil responses. The 78-item second section is concerned with both verbal and nonverbal expressions of affect (either negative or positive) by either teacher or students. (See Appendix I for a copy of the instrument.)

Research with the modified instrument (FLACCS) has continued to yield predictive associations with student outcome data (Soar, 1973). Reliability figures for five FLACCS factors in that study (Soar, 1973) were reported as .99, .89, .77, .99, and .98. Coefficients for 4 factors were not reported because of the lack of variability among teachers and the low incidence of occurrence of items in the factors. It was concluded, however, that the observers had attained significant reliability on the 4 factors, since the factors did discriminate significantly between differing programs of instruction involved in the study. Ragosta (1974) reanalyzed the data from the 1973 study, using separate series of observations rather than pooling observations as had been done in the original study. She reported overall reliabilities of .84, .81, .78, .74, .69, .70, .83, .64, and .84 for the nine FLACCS factors. Again, since the observers, observer-training, and observation procedures for the 1973 study were essentially the same as for the FLACCS data reported in the current study, no further reliability studies were made.

Aspy Rating Scales of Meaning, Genuineness, and Respect

Although these scales have been recently reformulated, they have been shown to fit the criteria proposed by Remmers (1963) as necessary for accepting rating scales for use in research. Specifically, they have been demonstrated to meet Remmers' criteria in the following ways;

1. Objectivity: Teams of raters have been trained to reach and maintain inter-rater reliabilities of .85 (Pearsons' r; p .01) on each scale. (Aspy & Roebuck, 1973, 1975; Aspy, Black, & Roebuck, 1972)
2. Reliability: The research teams are made up of individuals who maintain intra-rater(rate-rerate on the same tapes) reliabilities of .85. (Aspy, Black, & Roebuck, 1972; Aspy & Roebuck, 1973, 1975)
3. Sensitivity: The finest discriminations usually made of Meaning, Genuineness, and Respect are on the order of much, some, little, or none. Aspy's scales differentiate five levels of the phenomena from virtual absence to nearly omnipresence.
4. Validity: The relevance of the scales to a defined area of investigation and to relevant behavioral science constructs was demonstrated in

Chapters I and II. The data from the scale has proven to be covariate with categories of Flanders' Interaction Analysis and measures of the cognitive functioning of students, and to be independent of teacher sex, race, years of teaching experience, and the grade level on which the teacher instructs. (Roebuck & Aspy, 1974a, 1974b, Aspy & Roebuck, 1974)

5. Utility: The scales are efficient enough that they have been utilized in both the research studies cited and for in-service training of teachers. In the in-service training, as little as three hours training per scale results in 80% agreement in the use of the scales. (Aspy, 1972; Aspy & Roebuck, 1973 1975; Aspy, Black, & Roebuck, 1972)

In the current study, inter-rater reliabilities for the Aspy Rating Scales were calculated on the total data set. Each rater's scores on a scale for the two segments of each teacher tape were averaged by the computer to yield a mean score for that teacher for that tape. Thus, 150 scores (3 for each of the 50 teachers in the study) were reported for each rater for each scale. Table 4 displays the Pearson's *r* correlation coefficients for reliability of all combinations of two raters for each scale.

Because Pearson's *r* is a measure of the linear relationship of the ratings of any two raters rather than a measure of the discrepancy between any two scores, an index of agreement was also calculated. Each scale consists of 5 levels with possible ratings within levels at intervals of .1 (for example, 2.2, 2.3, 4.5, 4.6, etc.), thus making a 41-point scale. In calculating the Index of Agreement, acceptable agreement was set at ± .3. Thus the raters would be in agreement within half-a-level of each other. This level of agreement is similar to that expected on the original form of the scales (Carkhuff, 1969). The percent of agreement for all combinations of two raters and for all three raters is reported in Table 4.

Table 4

INTER-RATER RELIABILITY FOR ASPY SCALES

Index	Raters	Scale		
		Meaning	Genuineness	Respect
Pearson's *r* Correlation Coefficients	1 with 2	0.9961	0.9957	0.9953
	1 with 3	0.9986	0.9987	0.9984
	2 with 3	0.9988	0.9984	0.9984
Percent of Agreement	1 with 2	85.84	83.43	84.64
	1 with 3	96.08	94.88	94.28
	2 with 3	96.39	93.67	93.37
	All Three	85.54	83.43	84.64

Note. Number of Tapes Rated = 150.

Children's Manifest Anxiety Scale

The Children's form of the Manifest Anxiety Scale (Castaneda, McCandless, & Palermo, 1956) consists of (1) a 42-item Anxiety Scale adapted from Taylor's (1953) adult form and (2) an 11-item Lie Scale designed by the authors but modeled after Taylor's use of items from the MMPI for the same purpose of detecting the tendency of the subject to falsify his responses. The students respond "Yes" or "No" to each of the items on the test. Test and retest reliabilities were reported by the authors as ranging from .70

to .89 for the Anxiety Scale and .53 to .85 for the Lie Scale in a sample of 300 students. The Reliabilities were reported separately for sex by grade-level groups.

Dependence-Proneness Scale

The Dependence-Proneness Scale (Flanders, Anderson, & Amidon, 1961) is a 45-item forced-choice test to which the pupil responds by selecting "agree" or "disagree" for each item on the test. A reliability coefficient of .68 was reported for a sample of 1,280 eighth-grade students. The scale was found to differentiate significantly groups of students who (1) differ in sex, (2) are less likely to take an extreme position on an opinionnaire, (3) differ in their preferences for type of teacher, and (4) differ in achievement patterns according to the kind (direct/indirect) of teacher by whom they are taught.

How I See Myself Scale

The How I See Myself Scale, Elementary Form (Gordon, 1968) is a pupil self-report of self-concept. The scale consists of 40 bi-polar constructs in the form of two contrasting statements about self. Between the two poles of the construct is a 5-point scale which the pupil marks to indicate (1) which of the statements he feels is the way he is most of the time, and (2) how strongly he is that way.
The scale yields scores representing five factors or aspects of self-concept: (1) Teacher-School, (2) Physical Appearance,

(3) Interpersonal Adequacy, (4) Autonomy, (5) Academic Adequacy. Each factor-score represents the sum of nonexclusive items. That is, the points for an item may be included in the factor-score for more than one factor.

Test-Retest reliability ranges from .78 to .89 for total test scores and .72 to .82 for factor-scores. Relationships between the Inferred Self-Concept scores (Combs, Soper, & Courson, 1963) and the HISM indicate that the two approaches are measuring somewhat different things, as (1) the correlations between HISM factors and Inference Scores are low, ranging from .11 to .39, and (2) the pupils report themselves more positively than their self-concepts were inferred by the observers (Gordon, 1968).

Iowa Tests of Basic Skills

The Iowa Tests of Basic Skills (Lindquist & Hieronymus, 1956) provide an evaluation of the generalized intellectual skills involved in vocabulary mastery, reading comprehension, arithmetic computation and problem-solving, and spelling. The test is provided in three different forms for administration at the beginning, middle, or end of the year. Both grade-placement norms and percentile placement within grade are supplied for each of the three time-points. It is in general and accepted usage in educational settings.

Reliabilities for the Iowa Tests of Basic Skills range from .84 to .96 for major tests and .70 to .93 for sub-tests.

The composite reliability for the whole test ranges from .97 to .98 for different grades.

Junior Index of Motivation

The Junior Index of Motivation (Frymier, 1965) is an 80-item scale which yields a score representing the student's motivation for achievement. Only 50 of the items are scored and the relationship of the items to motivation is not obvious. For each item, the student records a sign and a number (+1, +2, -1, -2) to indicate agreement or disagreement and slight or strong support for the item statement. The scale differentiates between underachievers and overachievers and is related to teacher judgments of pupil motivation. It is correlated with raw achievement measures at +.37 to +.40.

Lorge-Thorndike Intelligence Test

The Lorge-Thorndike Intelligence Test (Lorge & Thorndike, 1962) is a general intelligence test which yields both a verbal and a nonverbal score. It was standardized on a population of 136,000 students in 22 states. The split-half reliabilities range from .88 to .94, depending on the level and the sub-test. Inter-correlations of the verbal and the nonverbal scores range from .54 to .70.

Social Reaction Inventory

The Social Reaction Inventory is an adaptation (by the staff of the Institute for Development of Human Resources) of the Internal-External Motivation Scale (Rotter, 1966). The I-E scale is a 29-item, forced-choice test, which includes six "filler" items intended to make somewhat more ambiguous the purpose of the test. As adapted by the Institute, the Social Reaction Inventory consists of 28 of the 29 items with slight changes in wording to lower the reading level of the scale so that it can be handled by elementary students. Item No. 3 on the I-E (which referred to the relationship between war and personal involvement in government) was dropped from the scale. Test-Retest reliability for the I-E scale has ranged from .61 to .83 across a one-month period in group administrations. Both construct and predictive validity at significant levels were reported in several studies summarized by the author (Rotter, 1966).

Torrance Tests of Creative Thinking

The Torrance Tests of Creative Thinking (Torrance, 1967) measure four aspects of creativity--fluency, flexibility, originality, and elaboration--with both verbal and figural scores for each aspect. In all sub-tests, the subjects give free responses to semi-structured tasks. The tests are scored by hand. _Fluency_ is the number of relevant responses given; _flexibility_ is the number of different categories of response; _originality_ is a sum of credits where some routine responses count zero, less common responses count for 1 point, and responses too infrequent to be on the list in the manual count 2 points; and _elaboration_ is the number of additional details per responses totaled over responses.

Test-retest reliabilities range from .50 to .93 over one to two-week periods and from .35 to .73 over three-year periods. The test manual summarizes more than 50 studies of the construct validity of the test but only one study of predictive validity involving criteria of real-life creative accomplishment.

Statistical Procedures

Since the purpose of this study has been to compare the relative effectiveness of the holistic versus the specific-behavior measures in predicting student gain, it was necessary to structure each set of measures as a set of predictors in a prediction model. This, in turn, necessitated the selection of a statistical model capable of using each of the respective sets of measures as a set of predictor variables. The regression model was selected as representing a predictive method appropriate for both cases.

Selection of regression as the predictive model of concern in the study then dictated the procedures involved in (1) the necessity for establishing independent data samples for each set of predictors, (2) the selection of predictor terms for each model, (3) the calculation of gain scores to serve as dependent variables, and (4) the selection of an appropriate test of the null hypotheses. The statistical procedures for splitting the sample, for the calculation of gain scores, for building and using the predictor models, and for testing the null hypotheses are delineated separately below.

Procedures for Splitting the Sample

In order to obtain the two desired independent samples for the statistical analyses, the students were stratified _within_ each classroom according to race-sex groups: black-male, white-male, black-female, and white-female. Then, one-half the students in each strata within each classroom were randomly selected and assigned to one of two sub-samples: i.e., they were assigned as a member of the sample whose outcome data would be used as the dependent variable for _either_ (1) the model using holistic variables as predictors _or_ (2) the model using specific-behavior variables as predictors. This stratified split-sample technique was chosen in order to eliminate possible undesirable effects of unequal race-sex distribution between the two independent samples being created.

Associated with each student was a data set which contained (1) his pretest and posttest data for all 16 educational outcome indices, (2) RCS and FLACCS measurements made in his classroom, (3) Aspy Rating Scale measurements made in his classroom, and (4) his I.Q. score. Thus, the procedure described above for identifying two sample _student_ sub-sets automatically created two independent sample _data_ sub-sets, _each_ of which would be used in the regression analyses for establishing the predictive power of one of the models.

Calculation of Gain Scores

The dependent variable in each regression analysis was the individual student's raw gain from pretest to posttest on one of the 16 educational outcome indices. Calculation of raw gain was simply the subtraction of pretest score from posttest score. In addition, the student's I.Q. score and his pretest score for the measure in which gain was the dependent variable were utilized as additional predictor variables in each regression; thus, an effect on gain was achieved similar to the process of adjusted gain, as in analysis of covariance.

Procedures for Building and Using the Prediction Models

Two regression models were built for each educational outcome index. In order to arrive at equivalent degrees of freedom, which tends to maximize the power of the subsequent tests of null hypotheses, it was necessary for each model to have the same number of terms.

Results of prior studies (Roebuck & Aspy, 1974b) indicated that for the Aspy Scales a nonlinear relationship might be most appropriate for use in the prediction of gain scores. Therefore, the predictors for the holistic variables model were established as (1) the quadratic expansion of the mean values of the Aspy Rating Scale scores for

Meaning, Genuineness, and Respect, plus (2) the teacher's individual standard deviation around her own mean for each of the three scales. This yielded 12 terms. Adding the student's I.Q. score and his pretest score, as two additional terms in the model, yielded a total of 14 predictors for the Holistic Model. The terms in the model are listed in Table 5.

Prior studies (Soar, 1971, 1973) indicated that the FLACCS and RCS factors have strong linear relationships to gain scores. Since it was desirable to equalize the number of terms in the two models, and since there were a total of 15 RCS and FLACCS factors from which terms to be used as predictors could be selected, it was determined that treating the specific-behavior measures in strictly linear fashion would be an acceptable representation of the predictive relationship. That is, the slight additional information which might be provided by using nonlinear relationships of variables that had already demonstrated strong linear relationships was evaluated as not efficient when considered against its two consequences: (1) reducing the power of the tests of hypotheses by unequal numbers of terms in the two models, or (2) reducing the number of RCS and FLACCS factors whose linear relationships could be considered in favor of using a quadratic expansion of the mean values of a smaller number of factors. Thus, 12 of the 15 RCS and FLACCS factors were selected for inclusion in linear fashion in the predictor

Table 5

PREDICTOR TERMS IN EACH OF THE REGRESSION MODELS

Model I: Holistic Variables		Model II: Specific-Behavior Variables	
Abbreviation	Name or Definition of Term	Abbreviation	Name or Definition of Term
P---*	Student's pretest score for measure used as dependent variable	P---*	Student's pretest score for measure used as dependent variable
I.Q.	Student's scores on I.Q. test	I.Q.	Student's score on I.Q. test
SMNG	Teacher's standard deviation around own mean for Meaning	FLC2	FLACCS Factor 2: Teacher Strong Control
SRES	Teacher's standard deviation around own mean for Respect	FLC3	FLACCS Factor 3: Teacher Positive Behavior
SGEN	Teacher's standard deviation around own mean for Genuineness	FLC5	FLACCS Factor 5: Pupil Supportive Behavior
MMNG	Mean value of Meaning scores	FLC6	FLACCS Factor 6: Teacher Positive Attention vs. Pupil Happy Response
MRES	Mean value of Respect scores	RCS1	RCS Factor 1: Teacher Initiation vs. Pupil Response
MGEN	Mean value of Genuineness scores	RCS2	RCS Factor 2: Student Talk vs. Teacher Direct/Indirect Behavior
MNG2	Squared value of Meaning scores	RCS3	RCS Factor 3: Teacher Acceptance/Flexibility
RXMG	Cross-product of Meaning and Respect	RCS4	RCS Factor 4: Teacher Acceptance/Rejection
RES2	Squared value of Respect scores	RCS5	RCS Factor 5: Extended Teacher Talk
GXMG	Cross-product of Meaning and Genuineness	RCS7	RCS Factor 7: Teacher Direction/Critic vs. Teacher Indirect Behavior
GXRS	Cross-product of Genuineness and Respect	RCS8	RCS Factor 8: Pupil Initiation vs. Teacher Response
GEN2	Squared value of Genuineness scores	RCS9	RCS Factor 9: Pupil Initiation vs. Pupil Response

*Blanks in abbreviation filled in by three letters in name of appropriate dependent variable.

model. Adding the student's I.Q. score and pretest score as additional terms, yielded a total of 14 terms for the specific-behaviors (RCS and FLACCS) model. Table 5 exhibits the terms used in the model.

Each of the predictor models thus derived was used to conduct 16 regression analyses, using, in turn, each of the 16 educational outcome indexes as the criterion (dependent variable). The 32 multiple R's (16 for each of the two models) resulting from these analyses thus constituted two independent sets of computed observations, which were then used in the tests of hypotheses.

Testing of Hypotheses

The multiple correlation coefficients (multiple R's) produced by the regression analysis procedures described above became the sample (data set) which was used for testing the null hypotheses. That is, the multiple R's obtained for each of the two predictor models for the same outcome index were compared to determine whether or not the amount of variance in gain accounted for by the Holistic Model was significantly higher than that accounted for by the Specific-Behaviors Model for that index. An appropriate test of the null hypothesis that the multiple R produced by the Holistic Model was no better than that produced by the Specific-Behaviors Model used Fisher's approximation (Anderson, 1958, pp. 78-79, Theorem 4.2.7). All tests were conducted at the .10 significance level.

The testing of the study hypothesis, therefore, involved the construction of 16 tests of null hypotheses associated with 16 outcome indices comparing multiple R's which had been calculated from two independent sets of data. Thus, each test was statistically legitimate. But because different measures from the same two data sets were used for each of the 16 tests, the tests were not statistically independent from each other; so, no generalized statistical conclusion could be drawn for the 16 tests as a complete set. However, a measurement for the overall predictive power of the holistic measures versus the specific-behavior measures could be represented by a count of the number of individual tests in which the alternative hypothesis was supported. Prior to the conduct of the tests, it was established that this counting procedure would be used to provide an estimate of the general predictive power of the holistic measures as opposed to that of the specific-behavior measures.

Bias of the Study

This study was biased in favor of the specific-behavior measures in that the predictive power of one set of holistic measures was matched against the predictive power of two sets of specific-behavior measures. That is, the predictive power of three scores from the Aspy Rating Scales was compared against the combined predictive power of 4 factor-scores from the Florida Climate and Control System (FLACCS) and 8 factor-scores from the Reciprocal Category System (RCS).

The extent of this bias is further evident when the item structure of the opposed sets of measures is examined. The Aspy Rating Scales yielded three items, each of which was one scale score. The FLACCS yielded tally-scores of 162 items which were subsequently combined into 87 item-scores and then factor-analyzed to yield 6 factors. (See Appendix IV for list of items and Appendix II for factor structure.) The RCS yielded a 20-by-20 matrix with a total of 400 cell-scores from which 61 variables were abstracted by selecting from column totals (sum of all cells in a column of the matrix), individual cell scores, or combinations of columns and/or cells. These 61 variables were then factor-analyzed to yield 9 factors. (See Appendix II for factor-structure and Appendix IV for the components of the 61 variables.) Thus, raw scores from the Aspy Rating Scales were compared against two sets of factor-scores which had been derived through actuarial manipulation of raw tally-scores from two instruments (RCS and FLACCS).

This 2-against-1 procedure, although biased against the holistic measures, was selected for the current study on the basis of the nature of the variables being considered. The Aspy Rating Scales tapped four dimensions of classroom process and considered the students' assignment of meanings to teacher behaviors in all four of those dimensions. The dimensions tapped were (1) instructional interactions, (2) expression of positive and negative affect, (3) levels of cognitive functioning, and (4) interpersonal relationships. In order to tease out the effects of the mediational elements of the paradigm (i.e., the consideration of student assignment of meanings to teaching acts) as opposed to the effects of the teaching acts themselves, it seemed desirable that the specific-behavior measures in the study account for behaviors in _all_ the dimensions tapped by the holistic measures.

No _one_ specific behavior instrument counted behaviors in all four dimensions of classroom process. The combination of the two selected instruments, however, did tap all four dimensions: FLACCS emphasized the measurement of interpersonal interactions and the expression of affect; the RCS accounted for instructional interactions and considered broad vs. narrow levels of cognitive functioning.

Therefore, it was decided to include measures from both the RCS and the FLACCS instruments in the Specific-Behaviors Model even though this would pose an extremely rigorous test for the Holistic Model. It was considered that even slight differences revealed by such a test would provide more information about the mediational elements of the paradigm than would larger differences from a less rigorous comparison. That is, if the holistic measures were compared against specific-behavior measures from only _one_ instrument, the source of positive results might be an effect of _either_ the mediational elements in the paradigm _or_ the additional dimensions tapped by the Aspy Scales but not tapped by the specific-behavior instrument. On the other hand, with all four of the dimensions tapped by the holistic measures also represented by specific-behavior measures (through the inclusion of two specific-behavior instruments), the source of any positive results could confidently be attributed to the mediational elements in the paradigm, i.e., the student's perceptions of the meaning of teaching acts.

CHAPTER IV

RESULTS OF ANALYSIS

The results of the analysis procedures described in Chapter III are presented in this chapter. The results are detailed separately for each of the three main steps of the procedures: (1) establishing two independent date sub-samples, (2) conducting regression analyses, and (3) testing of null hypotheses. Discussion of the results will be presented in the first part of Chapter V.

Means and Standard Deviations of Sub-Samples

As described in Chapter III, two independent data sub-samples were created by splitting the total sample. This was done for the purpose of obtaining statistical independence so that the subsequent tests of null hypotheses would be legitimate. The means and standard deviations for both gain scores and pretest scores for each of the outcome indices in each of the two sub-samples are presented in Tables 6 and 7. Table 6 presents the data for the sample used in regressions with holistic predictor terms, while Table 7 presents the data for the sample used in regressions with specific-behavior predictor terms.

Results of Regression Analyses

Two series of regression analyses were conducted. One

Table 6

MEANS AND STANDARD DEVIATIONS OF THE EDUCATIONAL OUTCOME INDICES FOR THE SAMPLE USED IN REGRESSION ANALYSES WITH HOLISTIC MEASURES AS PREDICTORS

Dependent Variable	Gain Parameters		Pretest Parameters	
	Mean	Standard Deviation	Mean	Standard Deviation
Anxiety Scale (Children's Manifest Anxiety Test)	-1.99	6.588	21.96	7.172
Arithmetic Test (Iowa Tests of Basic Skills)	6.07	8.208	41.93	9.094
Autonomy Self-Concept Factor (How I See Myself Scale)	-0.78	7.948	30.06	6.840
Dependence-Proneness Scale	0.63	4.939	26.63	4.588
Flexibility Factor (Torrance Tests of Creativity)	1.13	11.301	50.06	9.426
Fluidity Factor (Torrance Tests of Creativity)	6.58	10.358	45.34	7.879
Interpersonal Adequacy Factor (How I See Myself Scale)	1.21	13.013	63.36	12.346
Junior Index of Motivation	0.26	22.524	96.12	19.932
Lie Scale (Children's Manifest Anxiety Test)	-0.68	3.172	4.63	2.161
Originality Factor (Torrance Tests of Creativity)	6.44	10.919	45.84	6.992
Physical Appearance Factor (How I See Myself Scale)	0.16	7.224	27.383	6.558
Reading Test (Iowa Tests of Basic Skills)	3.81	8.363	43.56	11.828
Social Reaction Inventory	-0.51	3.192	10.37	2.637
Spelling Test (Iowa Tests of Basic Skills)	6.20	10.682	46.49	15.556
Teacher-School Factor (How I See Myself Scale)	-0.55	6.056	22.86	5.243
Vocabulary Test (Iowa Tests of Basic Skills)	6.70	9.265	40.56	11.861

Note. Number of Sample = 391.

Table 7

MEANS AND STANDARD DEVIATIONS OF THE EDUCATIONAL OUTCOME INDICES FOR THE SAMPLE USED IN REGRESSION ANALYSES WITH SPECIFIC-BEHAVIOR MEASURES AS PREDICTORS

Dependent Variable	Gain Parameters		Pretest Parameters	
	Mean	Standard Deviation	Mean	Standard Deviation
Anxiety Scale (Children's Manifest Anxiety Test)	-1.17	6.718	21.83	7.223
Arithmetic Test (Iowa Tests of Basic Skills)	5.82	9.206	42.26	9.577
Autonomy Self-Concept Factor (How I See Myself Scale)	-0.09	7.226	29.45	7.009
Dependence-Proneness Scale	0.13	4.895	26.41	4.679
Flexibility Factor (Torrance Tests of Creativity)	1.17	11.771	49.50	10.024
Fluidity Factor (Torrance Tests of Creativity)	5.47	10.273	45.34	8.312
Interpersonal Adequacy Factor (How I See Myself Scale)	-0.27	13.641	63.15	12.143
Junior Index of Motivation	-1.18	22.905	94.91	20.304
Lie Scale (Children's Manifest Anxiety Test)	-0.72	2.205	4.56	2.173
Originality Factor (Torrance Tests of Creativity)	5.14	10.735	45.70	6.969
Physical Appearance Factor (How I See Myself Scale)	-0.23	6.901	27.48	6.154
Reading Test (Iowa Tests of Basic Skills)	3.22	7.907	43.18	12.799
Social Reaction Inventory	-0.31	3.367	10.36	2.615
Spelling Test (Iowa Tests of Basic Skills)	6.59	10.738	45.64	16.023
Teacher-School Factor (How I See Myself Scale)	-0.71	6.323	22.12	5.388
Vocabulary Test (Iowa Tests of Basic Skills)	6.22	9.416	41.31	13.334

Note. Number of Sample = 390.

series examined holistic measures as predictors while the other series examined specific-behavior measures as predictors. Each series took each of the 16 outcome indices, in turn, as the dependent variable against which the appropriate predictor terms would be regressed.

Means and Standard Deviations of the predictor terms for the series of analyses using holistic measures are presented in Table 8. The regression equations derived for each of the outcome indices in this series of analyses are displayed in Exhibit I of Appendix III. Table 9 exhibits the results of the regression analyses for all 16 outcome variables. The multiple correlations achieved by the holistic measures ranged from .45 to .70 with all regressions significant at .0005 or better.

Means and Standard Deviations of the predictor terms for the series of analyses using specific-behavior measures are presented in Table 10. The regression equations derived in this series of analyses for each of the outcome indices is displayed in Exhibit II of Appendix III. Table 11 exhibits the results of the regression analyses for all 16 outcome variables. The multiple correlations achieved by the specific-behavior measures ranged from .47 to .65 with all regressions significant at .0005 or better.

Tests of Null Hypotheses

A test of the significance of the difference between the multiple R's achieved by each of the models for each of

Table 8

MEANS AND STANDARD DEVIATIONS OF THE PREDICTOR TERMS IN THE REGRESSION ANALYSES USING HOLISTIC MEASURES

Predictor Terms	Name or Description	Mean	Standard Deviation
P---	Student's pretest score for measure used as dependent variable	*	*
I.Q.	Student's scores on I.Q. test	94.9325	14.2706
SMNG	Teacher's standard deviation around own mean for Meaning	0.3461	0.1820
SRES	Teacher's standard deviation around own mean for Respect	0.2944	0.1373
SGEN	Teacher's standard deviation around own mean for Genuineness	0.2157	0.1306
MMNG	Mean value of Meaning scores	2.6909	0.3479
MRES	Mean value of Respect scores	2.0392	0.2959
MGEN	Mean value of Genuineness scores	2.4661	0.2459
MNG2	Squared value of Meaning scores	7.3623	1.8256
RXMG	Cross-product of Meaning and Respect	5.5812	1.4949
RES2	Squared value of Respect scores	4.2463	1.2717
GXMG	Cross-product of Meaning and Genuineness	6.7196	1.4668
GXRS	Cross-product of Genuineness and Respect	5.0957	1.2173
GEN2	Squared value of Genuineness scores	6.1429	1.1750

*See Table 6 for pretest parameters of outcome indices.

Table 9

RESULTS OF REGRESSION ANALYSES USING HOLISTIC VARIABLES AS PREDICTORS

Dependent Variable	Multiple Correlation	F-Value	Sig. Level
Gain on			
Anxiety Scale (Children's Manifest Anxiety Test)	0.45004	6.82106	.0005
Arithmetic Test (Iowa Tests of Basic Skills)	0.50819	9.35085	.0005
Autonomy Self-Concept Factor (How I See Myself Scale)	0.60973	15.89308	.0005
Dependence-Proneness Scale	0.51458	9.67298	.0005
Flexibility Factor (Torrance Tests of Creativity)	0.64004	18.63692	.0005
Fluidity Factor (Torrance Tests of Creativity)	0.56810	12.79836	.0005
Interpersonal Adequacy Factor (How I See Myself Scale)	0.69696	25.36909	.0005
Junior Index of Motivation	0.55951	12.23931	.0005
Lie Scale (Children's Manifest Anxiety Test)	0.47978	8.03101	.0005
Originality Factor (Torrance Tests of Creativity)	0.62430	17.15321	.0005
Physical Appearance Factor (How I See Myself Scale)	0.56970	12.90534	.0005
Reading Test (Iowa Tests of Basic Skills)	0.55474	11.93913	.0005
Social Reaction Inventory	0.59715	14.88452	.0005
Spelling Test (Iowa Tests of Basic Skills)	0.48890	8.43598	.0005
Teacher-School Factor (How I See Myself Scale)	0.57229	13.08009	.0005
Vocabulary Test (Iowa Tests of Basic Skills)	0.55291	11.82597	.0005

Table 10

MEANS AND STANDARD DEVIATIONS OF THE PREDICTOR TERMS IN REGRESSION ANALYSES USING SPECIFIC-BEHAVIOR MEASURES

Predictor Terms	Name or Description	Mean	Standard Deviation
P---	Student's pretest score for measure used as dependent variable	*	*
I.Q.	Student's score on I.Q. test	94.4246	15.2783
FLC2	FLACCS Factor 2: Teacher Strong Control	47.4678	5.5975
FLC3	FLACCS Factor 3: Teacher Positive Behavior	47.5297	4.1937
FLC5	FLACCS Factor 5: Pupil Supportive Behavior	49.6570	6.6014
FLC6	FLACCS Factor 6: Teacher Positive Attention vs. Pupil Happy Response	49.2713	6.8380
RCS1	RCS Factor 1: Teacher Initiation vs. Pupil Response	48.1040	5.7237
RCS2	RCS Factor 2: Student Talk vs. Teacher Direct/Indirect Talk	51.7543	8.0028
RCS3	RCS Factor 3: Teacher Acceptance/Flexibility	49.8117	6.2253
RCS4	RCS Factor 4: Teacher Acceptance/Rejection	50.9214	6.0402
RCS5	RCS Factor 5: Extended Teacher Talk	49.4321	7.0964
RCS7	RCS Factor 7: Teacher Direction/Criticism vs. Teacher Indirect Talk	48.5222	6.1740
RCS8	RCS Factor 8: Pupil Initiation vs. Teacher Response	53.1123	6.0663
RCS9	RCS Factor 9: Pupil Initiation vs. Pupil Response	53.3761	4.3183

*See Table 7 for pretest parameters of outcome indices.

Table 11

RESULTS OF REGRESSION ANALYSES USING SPECIFIC-BEHAVIOR VARIABLES AS PREDICTORS

Dependent Variable	Multiple Correlation	F-Value	Sig. Level
Gain on			
Anxiety Scale (Children's Manifest Anxiety Test)	0.47219	7.68611	.0005
Arithmetic Test (Iowa Tests of Basic Skills)	0.57765	13.41354	.0005
Autonomy Factor (How I See Myself Scale)	0.55477	11.90936	.0005
Dependence-Proneness Scale	0.55675	12.03247	.0005
Flexibility Factor (Torrance Tests of Creativity)	0.62817	17.45860	.0005
Fluidity Factor (Torrance Tests of Creativity)	0.53289	10.62278	.0005
Interpersonal Adequacy Factor (How I See Myself Scale)	0.64615	19.19865	.0005
Junior Index of Motivation	0.52452	10.16638	.0005
Lie Scale (Children's Manifest Anxiety Test)	0.59905	14.99261	.0005
Originality Factor (Torrance Tests of Creativity)	0.54049	11.05422	.0005
Physical Appearance Factor (How I See Myself Scale)	0.47831	7.94612	.0005
Reading Test (Iowa Tests of Basic Skills)	0.51425	9.63031	.0005
Social Reaction Inventory	0.62351	17.03670	.0005
Spelling Test (Iowa Tests of Basic Skills)	0.50236	9.04147	.0005
Teacher-School Factor (How I See Myself Scale)	0.58165	13.69538	.0005
Vocabulary Test (Iowa Tests of Basic Skills)	0.57917	13.52003	.0005

the outcome indices was conducted, using a one-tailed W statistic for Fisher's Approximation. Thus, 16 such comparisons were made, one for each of the 16 outcome indices. The results of the comparisons are displayed in Table 12.

The multiple R's produced by the prediction model using holistic measures were significantly larger than the multiple R's produced by the prediction models using specific-behavior measures in three of the 16 cases. The holistic prediction model did a significantly better job of predicting the variability in gain for Originality in Creativity, Interpersonal Adequacy Self-Concept, and Physical Appearance Self-Concept Factors.

The use of the one-tailed test as dictated by the form of the null hypotheses in the study does not detect any case in which the specific-behavior measures produced a multiple R significantly larger than that produced by the holistic measures. However, simply by reversing the sign of the test statistic, a test similar to the one used above could determine whether the Specific-Behaviors Model was better than the Holistic Model in any case. If this test had been applied to the data for this study, the results would have been that in only one case (Lie Scale of CMAS) was the multiple R produced by the Specific-Behavior Model larger than that produced by the Holistic Model at $p < .10$ significance. The significance and implications of these results will be discussed in Chapter V.

Table 12

TESTS OF THE HYPOTHESES USING ONE-TAILED FISHER'S APPROXIMATION W STATISTIC

Dependent Variable	Holistic Measures Model Multiple R's	Specific-Behaviors Model Multiple R's	W Statistic	Sig. Level
Gain on				
Anxiety Scale (Children's Manifest Anxiety Test)	.45000399	.47218996	-0.39162	
Arithmetic Test (Iowa Tests of Basic Skills)	.50818998	.57764995	-1.27304	
Autonomy Self-Concept Factor (How I See Myself Scale)	.60972995	.55476999	1.15872	
Dependence-Proneness Scale	.51457995	.55675000	-0.82363	
Flexibility Factor (Torrance Tests of Creativity)	.64003998	.62816995	0.27635	
Fluidity Factor (Torrance Tests of Creativity)	.56809998	.53288996	0.70349	
Interpersonal Adequacy Factor (How I See Myself Scale)	.69695997	.64614999	1.29037	.10
Junior Index of Motivation	.55950999	.52451998	0.68990	
Lie Scale (Children's Manifest Anxiety Test)	.47977996	.59904999	-2.35190	
Originality Factor (Torrance Tests of Creativity)	.62429994	.54048997	1.77016	.05
Physical Appearance Factor (How I See Myself Scale)	.56969994	.47830999	1.7585	.05
Reading Test (Iowa Tests of Basic Skills)	.55473995	.51424998	0.78939	
Social Reaction Inventory	.59714997	.62350994	-0.58491	

Table 10--continued.

Dependent Variable	Holistic Measures Model Multiple R's	Specific-Behaviors Model Multiple R's	W Statistic	Sig. Level
Spelling Test (Iowa Tests of Basic Skills)	.48889995	.50235999	-0.23838	
Teacher-School Factor (How I See Myself Scale)	.57228994	.58164996	-0.19531	
Vocabulary Test (Iowa Tests of Basic Skills)	.55290997	.57916999	-0.53798	

CHAPTER V

DISCUSSION AND IMPLICATIONS

Discussion

The results of this study support the hypothesis that holistic measures of classroom process which include the consideration of the student's assignment of meaning to teaching acts will do a better job of predicting student outcomes than will measures of classroom process which consider only the specific teaching acts themselves. That is, the study findings were that the Holistic Model did better than the Specific-Behavior Model in predicting 3 of the outcome indices, did as well in predicting 12 of the outcome indices, and did not do as well for only 1 index. The conclusion of the researcher was that, in fact, the Holistic Model did account for a larger amount of the variance among the educational outcome indices and that the study hypothesis was supported. The significance of this finding is enhanced when it is considered in relationship to the fact that holistic measures from only one instrument were compared with specific-behavior measures from two instruments.

Although the study procedure placed a burden of rigor on the holistic measures, it ensured that the differences obtained could confidently be attributed to the mediational elements of the holistic measures rather than to additional dimensions of classroom process tapped by them. That is, by inclusion of measures from two specific-behavior instruments, it became possible for the specific-behavior prediction model to include measures of behaviors in all the dimensions of classroom process considered by the holistic measures except that of the meaning assigned by students to teaching acts. Thus, the positive differences found in the study are not considered to be the effects of differences in the kinds of teaching acts which were measured but as the effects of the meanings attributed to those teaching acts by the students.

This necessity for using two specific-behavior instruments to cover all of the dimensions tapped by the one holistic instrument introduced a bias against the holistic measures, as the two specific-behavior instruments also registered behaviors in dimensions of classroom process (such as peer/peer interactions) which were not considered by the holistic measures. Therefore, the positive differences established in this study as the effects of the mediational elements of the paradigm (i.e., student attribution of meanings to teaching acts) are probably disproportionately smaller than the true effect.

A supplementary analysis of the data in the current study indicated that the predictive power of the holistic measures increased when students are classified as to race-sex characteristics. In this analysis, the predictive power of the holistic measures versus the specific-behavior measures was compared separately for each of the available race-sex sub-groups: Black-male, Black-female, White-male, and White-female. All procedures were the same as for the testing of the study hypotheses except that the data used in building the models were split-samples of students from only one race-sex group. Thus, 48 comparisons were made--16 for each sub-group of students. Of the 48 comparisons, 22 registered significant differences. Eighteen of the significant comparisons were in favor of the Holistic Model and 4 were in favor of the Specific-Behavior Model. All 4 of the differences in favor of the Specific-Behavior Model were found for male groups, with 2 occurring in each of the Black-male and White-male sub-groups. The 18 differences in favor of the Holistic Model were distributed among all 4 sub-groups with each sub-group being predicted more powerfully than the entire group had been in the major study.

Such an effect might have been anticipated from the work of those researchers who have demonstrated that groups of students with different characteristics have different perceptions of themselves and of the teaching-learning situation. Laird (1956) found that middle-class boys had different perceptions of themselves and of school than did lower-class boys. Hishiki (1969) established that Mexican-American girls reported their self-concepts and aspiration levels differently than did Southern white girls, and Davidson and Lang (1960) found both sex and social-class differences in students' perceptions of their teachers. Perhaps the most pertinent study along this line is Kallstrom's finding that

> Black female juveniles have a generally less positive view of self and perceive themselves as being less wanted, less worthy, less able, and less identified with others than do white female juveniles. These black female juveniles are also less open to experience than are the white female juveniles. (1970, p. 51)

He also found similar differences between delinquent and nondelinquent female juveniles (Kallstrom, 1970).

The establishment of effective predictive power for holistic measures related to the persons in the instructional process has important implications for both educational research and practice. These implications are discussed separately below.

Implications for Further Research

As a result of this study, several supplementary investigations seem indicated. Three specific studies are outlined below and a statement of general research implications is also made.

True Effects of the Mediational Elements

The first implication of this study for further research is related to the study bias. Although the current study established that measures which consider the students' attribution of meaning to teaching acts are effective in the prediction of educational outcomes, it did not establish the extent of the true effect of the mediational elements of those measures. That is, the inclusion in the specific-behaviors prediction model of measures of classroom process dimensions which were not considered in the holistic model was a contaminating factor which probably acted in such a fashion as to mask the true extent of the effects of the mediational elements considered in the holistic prediction model. Further studies need to be conducted to establish the true effect.

Such studies should seek closer matches of the kinds of behavior considered in the respective holistic and specific-behavior models to be compared. This, in turn, implies the need to develop additional holistic measures which will consider those dimensions of classroom process which are measured by the specific-behavior instruments but not now considered by available holistic instruments. Elements of classroom process for which such holistic instruments need to be developed include (1) student/student interaction, (2) student/task interaction, (3) student/goal interaction, and (4) teacher/goal interaction.

Differential Predictive Power of Holistic Model

The current study investigated the general predictive power of the holistic measures. There remains the possibility that there is a differential predictive relationship of the holistic measures with outcome variables. That is, it may be possible that the holistic measures may do a better job of predicting a particular type of outcome variable than they do of predicting another, different, type of outcome variable.

An examination of the pattern of prediction for the Holistic Model in this study supplies some support for this hypothesis. That is, if the multiple R's obtained by the holistic model (see Table 9) are arranged in order of size, the rank order from highest level of prediction to lowest level of prediction for the educational outcome variables is as follows:

1. Interpersonal Adequacy Self-Concept Factor
2. Flexibility Factor in Creativity
3. Originality Factor in Creativity
4. Autonomy Self-Concept Factor
5. Social Reaction Inventory
6. Teacher-School Self-Concept Factor
7. Physical Appearance Self-Concept Factor
8. Fluidity Factor in Creativity
9. Junior Index of Motivation
10. Reading Achievement
11. Vocabulary Achievement
12. Dependence-Proneness Scale
13. Arithmetic Achievement
14. Spelling Achievement
15. Lie Scale of the CMAS
16. Anxiety Scale of the CMAS

It seemed to this researcher that this rank order of prediction approximates a dimension of "more or less central to self." That is, it seemed that the arrangement of the outcome variables in terms of their rank order of prediction was similar to the order in which the outcome measures would be arranged if they were listed in terms of the degree to which they reflect the effects of mediation by perceptions lying at varying distances from the self within the phenomenal field. This hypothesis is currently being investigated.

Comparison with Perceptual Variables

A further study of the relative predictive power of ths holistic and specific-behavior models is suggested by a consideration of the nature of the variables in the two models. That is, the specific-behavior model only counted the behaviors occurring in the instructional situation. The holistic measures estimated the meaning of the sum of those behaviors and, in so doing, considered to some degree the student's assignment of meaning to teaching acts. However, the holistic measures were not strictly inferential. That is, they observed the teacher's behaviors as well as made inferences of the meanings assigned by students.

A further study should be made which would include a comparison with purely inferential measures. The rationale would be that since holistic measures which considered to some degree the assignment of meaning to behaviors did a better job of predicting subsequent educational outcomes than did the specific-behavior measures, then perceptualistic measures making direct inferences of the student's perceptions of his relationship to the teaching-learning situation might do an even better job of predicting subsequent outcomes.

General Research Implications

Although there remain to be made several specific studies of theoretical aspects or practical applications of the holistic measures, the current study has adequately established the utility of the holistic instrument as a research tool. This establishment of a presage-related process measure (i.e., a process measure directly related to the persons in the process) opens the door to a whole range of investigations of the expanded teacher effectiveness paradigm proposed in Figure 2 of Chapter I. As those studies are completed and as additional presage-related process measures are developed and validated, it should be possible to supply the data now "lacking on all fronts" as Soar (1968) put it.

Further general implications for educational research arise from the economical procedures involved in applying the holistic instrument. First, data analysis with a holistic measure is conducted with raw means. Thus, there are no intervening computational processes of (1) calculating item-scores from behavioral counts (tallies) and then (2) conducting factor-analysis. This amounts to a considerable savings in data processing time and cost. It also allows the establishment of much more direct relationships of the process variables to the criterion variables. That is, there is no intervening interpretative step such as becomes involved in the characterizing and naming of factors.

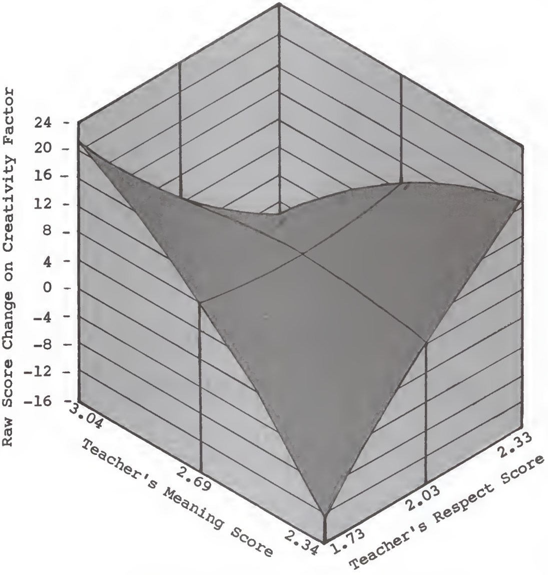
Second, in this study at least, data-collection for the holistic measures was more economical than for the specific-behavior measures. One set of raters using one holistic instrument provided data which did a more effective job of predicting outcome measures than the data provided by two sets of raters using two specific-behavior instruments.

A last implication is derived from the smaller number of variables involved in the holistic instrument. Because the instrument has fewer variables (essentially only three scale-scores from which the quadratic expansions and standard deviations are calculated), it is possible to demonstrate the dynamic interaction of all the holistic variables with the criterion variables through the process of Response Surface Analysis. That is, if one substitutes a set of three data points (mean, plus one standard deviation and minus one standard deviation) for each of the scales into the equation obtained through regression analysis, it is possible to generate a set of plot points which allow the visualization of the interaction of the variables. Figures 3 and 4 are examples of such response surfaces which were generated from the regression equations obtained in this study.

The value of this procedure for research purposes is perhaps best visualized by a demonstration of the kinds of information provided by the surfaces. In Figures 3 and 4, Genuineness is held constant at its Mean value for the sample while the values for Respect and Meaning are varied. The dependent (response) variable for Figure 4 is gain on Reading achievement and for Figure 3 is gain on Originality factor in Creativity. The values for the response variable are read from height of the surface above the base of the cube.

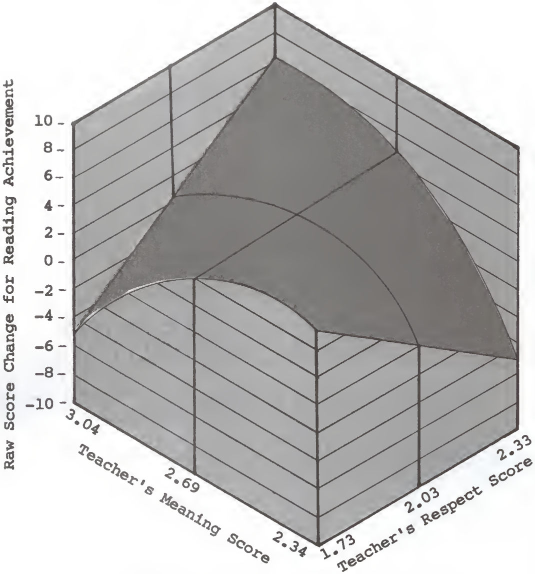
The particular contribution of response surface procedure to educational research is demonstrated in Figure 3 by a comparison of the information yielded by the lines on the front surfaces of the cube in contrast to the information yielded by the lines on the rear surfaces of the cube. The front surfaces indicate, as would be anticipated, that when _either_ the Teacher's Respect or Meaning increases, so does student gain on Originality. The rear surfaces indicate, however, that the interaction of Respect and Meaning with Originality is curvilinear. Therefore, at high levels of _both_ Meaning and Respect, student gain is decreased.

Figure 4 shows the response surface for student gain in Reading achievement. It demonstrates that the Teacher's Respect has an almost strictly linear relationship with Reading achievement but that the direction of this relationship is determined by the Teacher's level of Meaning. That is, at low levels of Meaning, gain in Reading is _negatively_ correlated with increasing Respect; but at high levels of Meaning, gain in Reading is _positively_ correlated with increasing Respect.



Teacher's Genuineness Score = 2.46

Figure 3. Response Surface for Change in Originality Factor of Creativity



Teacher's Genuineness Score = 2.46

Figure 4. Response Surface for Change in Reading Achievement

The particular value of the holistic measures as predictors for use in Response Surface Analysis is that a set of three such figures suffices to visualize the interaction of all the holistic measures with an outcome variable. In comparison, to accomplish this same visualization of the interaction of all predictor variables with an outcome variable would take 729 surfaces for the 8 factors of the RCS instrument used in this study.

Implications for Educational Practice

This study had several implications for educational practice. These implications for teacher training and for accountability are discussed separately below.

Teacher Training

The first implication of this study for teacher training--either in-service or pre-service--is drawn from the conclusion that the mediational elements of the paradigm were shown to have an important predictive relationship with educational outcomes. That is, it was shown that consideration of the meanings a student assigns to teaching acts is an important part of the relationship of those acts with subsequent educational outcomes. The implication, then, is that teachers need (1) to know the importance of student perceptions of classroom process, (2) to learn to infer how the student is perceiving the learning situation, and (3) to take those inferred students' perceptions into account in setting instructional goals and in carrying out instructional activities.

This implication, in itself, further implies that teacher education programs need to take the directions recommended by those educators (Combs, 1965, 1972; Rogers, 1969; Lindgren, 1954; Jersild, 1965; Dandes, 1966; Del Popolo, 1965) who have continued to maintain that teacher selection and training must be concerned with "who" the teacher is, i.e., with her perceptual organization and her skills in assessing the perceptual organization of her students. The New Elementary Program at the University of Florida (Avila et al., 1972; Blume, 1971; Wass et al., 1974) is an example of the kind of teacher training program which is supported by the results of this study.

A second implication of the study arises from the nature of the holistic measures used. Since they are few in number and the data used in analysis are the raw means, it is possible to communicate the results directly to teachers. That is, the major relationships of the variables to research results are direct and self-evident. They can be understood by persons with relatively little statistical training (e.g., without knowledge of T-scores, factor-analysis, etc.) particularly when the results are presented in visual form as in the response surfaces discussed previously.

Although information in itself will not serve to change behavior (Combs, 1972), information plus appropriate goals and an available technology may. This leads directly to a third implication. The holistic measures are appropriate for teacher training and should be utilized. That is, the teacher who wants to change in the directions implied by the research can learn to apply the scales to his own teaching behavior. A further application of the scales for this purpose would be the development of a technology such that a teacher could receive feedback as to his class's predicted individual placement on the response surface for an outcome of interest to him. That is, given that a large enough data base (several hundred classrooms) had been established so that a generalized prediction formula had been developed for each of the outcome criteria of interest, a teacher's data could be plugged into the formula and his class's predicted outcome position on the response surface calculated and returned in visual form. He could then determine the direction he wanted to move. The coordinates of the surface for the point he determined as his outcome goal would suggest appropriate skill goals for his consideration and/or adoption.

Accountability

Combs (1972) pointed out the need for accountability assessments to go beyond behavioral objectives and to incorporate goals and assessment procedures which lend themselves to the "effective production of intelligent behavior as a primary goal of modern education." (Combs, 1972, p. 12) He went on to assert that such a process calls for holistic approaches to the assessment of both student learning and teaching behavior.

The current researcher believes that this study supports Combs' viewpoint in that it indicates (1) that holistic measures can be developed which are at least as effective in educational measurement as the specific-behavior measures which are currently in widespread use and (2) that holistic measure related to the persons in the process will be especially capable of predicting and/or assessing those kinds of student behaviors which are related to the production of intelligent behavior. This last implication is based on the rank order of the prediction of educational outcomes as presented earlier in this chapter. That is, the educational outcomes which were most highly predicted by the holistic measures in this study were those outcomes--such as creativity, interpersonal adequacy, and autonomy--which might be expected to be related to intelligent problem-solving behavior.

The implication for educational practitioners is obvious. As recommended by Combs (1972), educators need to devote time, energy, and money to devising ways to implement accountability procedures related to humanistic goals and methods. This study implies that it can be done. The question is only will it be done?

Summary

This study compared the efficiency of holistic measures of classroom process against specific-behavior measures of classroom process in terms of how well they could predict student change on educational outcome indices. The holistic instrument used in the study was the Aspy Rating Scales of Meaning, Genuineness, and Respect. The specific-behavior instruments were the Reciprocal Category System and the Florida Climate and Control System. Educational outcome indices were 16 paper and pencil measures ranging from academic achievement tests through self-reports of self-concept.

The sample was sub-divided into two independent data sub-samples for the purpose of the statistical analysis. Each sub-sample provided the data-set for regression analyses conducted with one of the two models. The multiple R's produced by regression analyses using holistic measures as predictor terms were compared against the multiple R's produced by regression analyses using specific-behavior measures as predictor terms. Null hypotheses were tested with the W statistic for Fisher's Approximation. It was found that the holistic measures did a significantly better job of predicting three of the outcome indices, performed as well for 12 of the outcome indices, and did not do as well as the specific-behavior measures for only 1 index. Thus, it was concluded that the study hypothesis of greater predictive power for the holistic measures was supported.

Supplementary analyses of the data indicated that the predictive power of the holistic measures increased when students were sampled from race-sex groups. Several specific studies arising from the current study were outlined and general research implications were stated. Implications for educational practice in the areas of teacher-training and accountability were discussed.

REFERENCES

Amidon, E., & Flanders, N. A. The effects of direct and indirect teacher influence on dependent-prone students learning geometry. Journal of Educational Psychology, 1961, 52, 286-291.

Amidon, E. J., & Hough, J. B. Interaction analysis: Theory, research and application. Reading, Mass.: Addison-Wesley, 1967.

Anderson, H. H. The measurement of domination and of socially integrative behavior in teachers' contacts with children. Child Development, 1939, 10, 73-89.

Anderson, H. H. Studies of teachers' classroom personalities, III. Stanford, Calif.: Stanford University Press, 1946.

Anderson, T. W. Introduction to multivariate statistical analysis. New York: John Wiley and Sons, 1958.

Anthony, B. C. M. The identification and measurement of classroom environmental process variables related to academic achievement. Doctoral dissertation, University of Chicago, 1967.

Aspy, D. N. Counseling and education. In R. R. Carkhuff (Ed.,), The counselor's contribution to facilitative processes. Urbana, Ill.: Parkinson, 1967.

Aspy, D. N. The effect of teacher-offered conditions of empathy, positive regard, and congruence upon student achievement. Florida Journal of Educational Research, 1969, 2, 39-48.

Aspy, D. N. The effect of the teacher's inferred self-concept upon student achievement. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, February 1968.

Aspy, D. N. Toward a technology for humanizing education. Champaign, Illinois: Research Press, 1972.

Aspy, D. N., Black, B., & Roebuck, F. N. The relationship of teacher-offered conditions of respect to behaviors described by Flanders' Interation Analysis. Journal of Negro Education, 1972, 61, 370-376.

Aspy, D. N., & Hadlock, W. The effect of empathy, warmth, and genuineness on elementary students' reading achievement. Reviewed in Truax, C. B., & Carkhuff, R. R. Toward effective counseling and psychotherapy. Chicago: Aldine Publishing Company, 1967.

Aspy, D. N., & Roebuck, F. N. An investigation of the relationship between student levels of cognitive functioning and the teachers' classroom behavior. Journal of Educational Research, 1973, 65, 365-368.

Aspy, D. N., & Roebuck, F. N. The relationship of teacher-offered conditions of meaning to behaviors described by Flanders' Interaction Analysis. Educational Quarterly, 1975, 95, 216-222.

Aspy, D. N., & Roebuck, F. N. Research summary: Effects of training in interpersonal skills. Interim report No. 4 for NIMH Grant No. 5 PO 1 MH 19871. Monroe: Northeast Louisiana University, 1974.

Avila, D. L., Combs, W., Olson, W., Packer, A., & Shea, J. The Florida experimental program in elementary education. Improving College and University Teaching. Spring 1972, 148-149.

Benton, J. A. Perceptual characteristics of Episcopal pastors. Doctoral dissertation, University of Florida, 1964.

Bledsoe, J. C., Brown, D., & Strickland, A. D. Factors related to pupil observation reports of teachers and attitudes toward their teacher. Journal of Educational Research, 1971, 65, 119-126.

Blume, R. Humanizing teacher education. Phi Delta Kappan, 1971, 52, 411-415.

Brown, R. G. A study of the perceptual organization of elementary and secondary Outstanding Young Educators. Doctoral dissertation, University of Florida, 1970.

Campbell, J. R. Cognitive and affective process development and its relation to a teacher's interaction ratio. Journal of Research in Science Teaching, 1971, 8, 317-323.

Carkhuff, R. R. Helping and human relations (2 vols.). New York: Holt, Rinehart and Winston, 1970.

Carkhuff, R. R., & Berenson, B. G. Beyond counseling and therapy. New York: Holt, Rinehart and Winston, 1967.

Castaneda, A., McCandless, B. R., & Palermo, D. S. The children's form of the manifest anxiety scale. Child Development, 1956, 27, 317-323.

Choy C. The relationship of college teacher effectiveness to conceptual systems orientation and perceptual orientation. Doctoral dissertation, University of Northern Colorado, 1969.

Cogan, M. L. The behavior of teachers and the productive behavior of their pupils. Journal of Experimental Education, 1958, 27, 89-124.

Combs, A. W. The professional education of teachers. Boston: Allyn and Bacon, 1965.

Combs, A. W. Educational accountability: Beyond behavioral objectives. Washington, D. C.: Association for Supervision and Curriculum Development, 1972.

Combs, A. W., & Snygg, D. Individual behavior. New York Harper, 1959.

Combs, A. W., & Soper, D. W. The perceptual organization of effective counselors. Journal of Consulting Psychology, 1963, 10, 222-226.

Combs, A. W., Soper, D. W., & Courson, C. C. The measurement of self concept and self report. Educational and Psychological Measurement, 1963, 23, 493-500.

Combs, A. W., Soper, D. W., Gooding, C. T., Benton, J. A., Jr., Dickman, J. F., & Usher, R. H. Florida studies in the helping professions. University of Florida Monographs, 1969, 37.

Cook, W. W., Kearney, N. C., Rocchio, P. D., & Thompson, A. Significant factors in teachers' classroom attitudes. Journal of Teacher Education, 1956, 7, 274-279.

Cooper, J. G., & Lewis, R. B. Qualitative Rorschach factors in the evaluation of teacher effectiveness. Journal of Educational Research, 1951, 44, 703-707.

Courson, C. The use of inference as a research tool. Educational and Psychological Measurement, 1965, 25, 1029-1038.

Dandes, H. M. Psychological health and teaching effectiveness. Journal of Teacher Education, 1966, 17, 301-306.

Davidson, H. H., & Lang, G. Children's perceptions of their teachers' feelings toward them related to self-perception, school achievement and behavior. Journal of Experimental Education, 1960, 29, 107-118.

Davies, L. S. Some relationships between attitudes, personality characteristics, and verbal behavior of selected teachers. Doctoral dissertation, University of Minnesota, 1961.

Dellow, D. W. A study of the perceptual organization of teachers and conditions of empathy, congruence, and positive regard. Doctoral dissertation, University of Florida, 1971.

Del Popolo, J. A. Teacher personality: A concern of teacher education. Peabody Journal of Education, 1965, 43, 50-55.

Dickman, J. F. The perceptual organization of person-oriented versus task-oriented student nurses. Doctoral dissertation, University of Florida, 1967.

Dixon, W., & Morse, C. The prediction of teaching performance: Empathic potential. Journal of Teacher Education, 1961, 12, 322-329.

Doyle, E. J. The relationship between college teacher effectiveness and inferred characteristics of the adequate personality. Doctoral dissertation, University of Northern Colorado, 1969.

Emmerling, F. C., Jr. A study of the relationships between personality characteristics of classroom teachers and pupil perceptions of these teachers. Doctoral dissertation, Auburn University, 1961.

Flanders, N. A. Teacher influence, pupil attitudes and achievement. (U. S. Department of Health, Education and Welfare, Cooperative Research Monograph No. 12). Washington, D. C.: U. S. Government Printing Office, 1965.

Flanders, N. A. Teacher effectiveness. In R. L. Evel (Ed.), Encyclopedia of educational research (4th ed.). Toronto: Macmillan, 1969.

Flanders, N. A. Analyzing teaching behavior. Reading, Mass.: Addison-Wesley, 1970.

Flanders, N. A. Teacher influence in the classroom. In E. J. Amidon, & J. B. Hough (Eds.), Interaction analysis: Theory, research, and application. Reading, Mass.: Addison-Wesley, 1967.

Flanders, N. A., Anderson, J. P., & Amidon, E. J. Measuring dependence-proneness in the classroom. Educational and Psychological Measurement, 1961, 21, 575-587.

Fowler, B. D. Relations of teacher personality characteristics and attitudes to teacher-pupil rapport and emotional climate in the elementary classroom. Doctoral dissertation, University of South Carolina, 1962.

Fox, D. J. The research process in education. New York: Holt, Rinehart and Winston, 1969.

French, L. Individualizing classroom communication. Educational Leadership, 1970, 28, 193-196.

Frymier, J. R. Development and validation of a motivation index: A sixth report. Columbus: Ohio State University, 1965.

Furst, N. F. The multiple languages of the classroom: A further analysis and a synthesis of meanings communicated in high school teaching. Doctoral dissertation, Temple University, 1967.

Gage, N. L. Desirable behaviors of teachers. Urban Education, 1965, 1, 85-95.

Gage, N. L., Leavitt, G. S., & Stone, G. C. Teachers' understanding of their pupils and pupils' ratings of their teachers. Psychological Monographs, 1955, 69 (2, Whole No. 406).

Getzels, J. W., & Jackson, P. W. The teacher's personality and characteristics. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

Gillett, L. R. Affective relationships among student teachers' acceptance of self, acceptance of pupils, and acceptance by pupils. Doctoral dissertation, University of North Dakota, 1965.

Gooding, C. T. An observational analysis of the perceptual organization of effective teachers. Doctoral dissertation, University of Florida, 1964.

Gordon, I. J. A test manual for the How I See Myself Scale. Gainesville: Florida Educational Research and Development Council, 1968.

Gray, F., Graubard, S., & Rosenberg, H. Little brother is changing you. Psychology Today, March 1974, 42-46.

Guba, E. G., & Getzels, J. W. Personality and teacher effectiveness: A problem in theoretical research. Journal of Educational Psychology, 1955, 46, 330-344.

Harris, C. W. (Ed.). Problems in measuring change. Madison: University of Wisconsin Press, 1963.

Harris, A. J. A continuation of the CRAFT project: Comparing reading approaches with disadvantaged urban Negro children in primary grades. U. S. Office of Education Project No. 5-0570-2-12-1. New York: City University of New York, 1968.

Harvey, D. J. Teachers' beliefs, classroom atmosphere and student behavior, final report. Washington, D. C.: Office of Economic Opportunity, 1967. (ERIC Document Reproduction Service No. ED 018249)

Hays, E. R. Ego-threatening classroom communication: A factor analysis of student perceptions. Speech Teacher, 1970, 19, 43-48.

Heil, L. M. Modifying behaviors (self-concept) of certain prospective teachers. Washington, D. C.: Government Printing Office, 1962. (ERIC Document Reproduction Service No. ED 003 050)

Heil, L. M., & Washburne, C. Brooklyn College research in teacher effectiveness. Journal of Educational Research, 1962, 55, 346-351.

Hishiki, P. C. The self concepts of sixth grade girls of Mexican-American descent. California Journal of Educational Research, 1969, 20, 56-62.

Horn, E. Distribution of opportunity for participation among the various pupils in classroom recitations. Teacher's College Contributions to Education, No. 67. New York: Columbia University Press, 1914.

Hughes, M. Development of the means for the assessment of the quality of teaching in elementary schools. Salt Lake City: University of Utah Press, 1959.

Hunter, C. P. Classroom climate and pupil characteristics in special classes for the educationally handicapped. Doctoral dissertation, University of Southern California at Los Angeles, 1969.

Jersild, A. T. The voice of the self. National Education Association Journal, 1965, 54, 23-25.

Kallstrom, D. W. A study of the perceptual organization of black and white delinquent and non-delinquent juvenile females. Doctoral dissertation, University of Florida, 1970.

Laird, D. C. S. How eleven-year-old boys see their teachers. Progressive Education, 1956, 5, 115-118.

Lamke, T. A. Personality and teaching success. Journal of Experimental Education, 1951, 20, 217-259.

LaShier, W. S., & Westmeyer, P. The use of interaction analysis in BSCS laboratory block classrooms. Journal of Teacher Education, 1967, 13, 439-446.

Leeds, C. H. Teacher attitude and temperament as a measure of teacher-pupil rapport. Journal of Applied Psychology, 1956, 40, 333-337.

Lessinger, L., & Allen, D. Performance proposals for educational funding: A new approach to federal resource allocation. Phi Delta Kappan, 1969, 51, 136-137.

Levy, L. H. The conditioning and generalization of changes in social perceptual dispostions. Journal of Abnormal and Social Psychology, 1961, 63, 583-587.

Lindgren, H. C. Mental health in education. New York: Henry Holt, 1954.

Lindquist, E. F., & Hieronymus, A. N. Iowa tests of basic skills. New York: Houghton Mifflin, 1956.

Lippitt, R., & White, R. K. The "social climate" of children's groups. In R. R. Barker, J. S. Kounin, & H. F. Wright (Eds.), Child behavior and development. New York: McGraw-Hill, 1943.

Lipton, A. Relationship of teacher rigidity to progress of retarded readers: A study of the relationship between teachers' cognitive and social rigidity patterns and the reading achievement gains of retarded readers. Dissertation Abstracts, 1967, 27, 4170A.

Lorge, I., & Thorndike, R. L. Technical manual for Lorge-Thorndike intelligence tests. New York: Houghton-Mifflin, 1962.

Malpass, L. F. Some relationships between students' perceptions of school and their achievement. Journal of Educational Psychology, 1953, 44, 475-482.

Medley, D. M., & Mitzel, H. E. A technique for measuring classroom behavior. Journal of Educational Psychology, 1958, 49, 86-92.

Medley, D. M., & Mitzel, H. E. Measuring classroom behavior by systematic observation. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

Moore, C. H., & Cole, D. The relation of MMPI scores to practice teaching ratings. Journal of Educational Research, 1957, 50, 711-715.

Morgenstern, R. E. The relationship between two modes of interpersonal conditions and college teacher effectiveness. Doctoral dissertation, University of Northern Colorado, 1969.

Morrison, B. M. The reactions of internal and external children to patterns of teaching behavior. Doctoral dissertation, University of Michigan, 1966.

Morsh, J. E., & Wilder, E. W. Identifying the effective instructor: A review of quantitative studies, 1900-1952. USAF Research Bulletin No. AFPTRC-TR-54-44. San Antonio, Tex.: Personnel Training Research Center, 1954.

Myers, R. E. A comparison of the perceptions of elementary school children in open area and self-contained classrooms in British Columbia. Journal of Research and Development in Education, 1971, 4, 100-106.

Ober, R. L., Wood, S. E., & Roberts, A. The development of a reciprocal category system for assessing teacher-student verbal interaction. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, February 1968.

Pendergrass, P. W. The relationship of perceived threat to adolescent students' feelings of adequacy in the classroom. Doctoral dissertation, University of Florida, 1971.

Perkins, H. V. Climate influences group learning. Journal of Educational Research, 1951, 45, 115-119.

Perkins, H. V. Classroom behavior and under-achievement. American Educational Research Journal, 1965, 2, 1-14.

Pierce, R. J. An investigation of grade-point average and therapeutic process variables. Doctoral dissertation, University of Massachusetts, 1966.

Puckett, R. C. Making supervision objective. School Review, 1928, 36, 209-212.

Ragosta, M. Reliability in classroom observation, observer effects and stability of behavior. Doctoral dissertation, University of Florida, 1974.

Remmers, H. H. Rating methods in research on teaching. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

Roebuck, F. N., & Aspy, D. N. Grade-level contributions to variance of Flanders' interaction categories. Journal of Experimental Education, 1974, 42, 86-92. (a)

Roebuck, F. N., & Aspy, D. N. Response surface analysis. Interim report No. 3 for NIMH Grant No. 5 PO 1 MH 19871. Monroe: Northeast Louisiana University, 1974. (b)

Rogers, C. R. On becoming a person. Boston: Houghton-Mifflin, 1961.

Rogers, C. R. Freedom to learn. Columbus, Ohio: Merrill, 1969.

Rosenshine, B., & Furst, N. The use of direct observation to study teaching. In R. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally, 1973.

Rotter, J. B. Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 1966, 80 (1, Whole No. 609).

Ryan, K. A. A plan for a new type of professional training for a new type of teaching staff. National Commission on Teacher Education, Occasional Paper No. 2, 1968.

Ryans, D. G. Assessment of teacher behavior and instruction. Review of Educational Research, 1963, 33, 415-441.

Sarason, S. B. A study of teacher behavior in relation to children differing in anxiety level. Washington, D. C.: Cooperative Research Program, 1962. (ERIC Document Reproduction Service No. ED 002 911)

Schmuck, R. Some aspects of classroom social climate. Psychology in the Schools, 1966, 3, 59-65.

Silberman, M. L. Behavioral expression of teachers' attitudes toward elementary school students. Journal of Educational Psychology, 1969, 60, 402-407.

Smith, L. M., & Kleine, P. F. Teacher awareness: Social cognition in the classroom. The School Review, 1969, 245-256.

Soar, R. S. An integrative approach to classroom learning. Final Report, Public Health Service Grant No. 5-R11-MH 01096, & NIMH Grant No. 7-R11-MH 02045. Philadelphia: Temple University, 1966. (ERIC Document Reproduction Service No. ED 033-479) (a)

Soar, R. S. Research in the teaching processes: Four studies with the OSCAR technique. Paper delivered to the National Society of College Teachers of Education, 1966. (b)

Soar, R. S. The study of presage-process-production relationships: Implications for classroom process measurement. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, February 1968.

Soar, R. S. Follow through classroom process measurement, Gainesville, Florida. Research report, contract PEG-O-8-522471-4618 (100) & OEG 0-8-522394-3991 (286). Gainesville: University of Florida, 1971.

Soar, R. S. Final report: Follow through classroom process and pupil growth. Report, contract OEG-O-8-522394-3991 (286). Gainesville: University of Florida, 1973.

Soar, R. S., Soar, R. M., & Ragosta, M. H. Florida Climate and Control System: Training Manual. Gainesville: University of Florida, 1971.

Sorensen, P. Program for learning in accordance with needs. Phi Delta Kappan, November, 1970, 180-181.

Taylor, J. A. A personality scale of manifest anxiety. Journal of Abnormal and Social Psychology, 1953, 48, 285-290.

Thatcher, D. A. Ask not what you can do for the machine . . . . Phi Delta Kappan, April, 1972, 509-511.

Thomas, D. S. Some new techniques for studying social behavior. Child Development Monographs, 1929. (1)

Torrance, E. P. Torrance Tests of Creative Thinking: Administrators' Manual. Princeton, N. J.: Personnel Press, 1967.

Truax, C. B., & Carkhuff, R. R. Toward effective counseling and psychotherapy. Chicago: Aldine, 1967.

Truax, C. B., & Tatum, C. R. An extension from the effective psychotherapeutic model to constructive personality change to preschool children. Childhood Education, 1966, 42, 456-462.

Tuckman, B. W. A technique for the assessment of teacher directiveness. Journal of Educational Research, 1970, 63, 395-400.

Usher, R. H. The relationship of perceptions of self, others, and the helping task to certain measures of college faculty effectiveness. Doctoral dissertation, University of Florida, 1966.

Vonk, H. G. The relationship of teacher effectiveness to perception of self and teaching purposes. Doctoral dissertation, University of Florida, 1970.

Wallen, N. E. Relationships between teacher characteristics and student behavior--part III. Washington, D. C.: U. S. Government Printing Office, 1966. (ERIC Document Reproduction Service No. ED 010 390)

Wass, H., Blume, R. A., Combs, A. W., & Hedges, W. D. Humanistic teacher education: An experiment in systematic curriculum innovation. Fort Collins, Colo.: Shields, 1974.

Willard, R. A. A study of the relationships between the valued behaviors of selected teachers and the learning experience provided in their classrooms. Journal of Educational Research, 1955, 49, 45-51.

Withall, J. The development of a climate index. Journal of Educational Research, 1951, 45, 93-99.

Wright, B., & Sherman, B. Love and mastery in the child's image of the teacher. The School Review, Summer 1965, 89-101.

APPENDIX I

Instruments for Measuring Classroom Process

Exhibit 1

ASPY RATING SCALES

Understanding of Meaning

A Scale for Measurement*

Level 5	The tone and words of the teacher's verbal communication always add to the students' meanings. The students are always helped to express deeper meanings of their experiences.
Level 4	The tone and the words of the teacher's verbal communication consistently add to the students' meanings. The students are usually helped to express deeper meanings of their experiences.
Level 3	The tone and the words of the teacher's verbal communication match the students'. That is, they neither detract from nor add to the students' expression of meaning.
Level 2	The tone of the teacher's verbal communication indicates she understands the most obvious meanings of the student's experience to him. However, the teacher's verbal communications are less intense than the student's. That is, she detracts from his expression of meaning.
Level 1	The teacher's verbal communication indicates no response to the student's feeling. The teacher seems totally unaware of the meaning of the experience to the student.

*David N. Aspy

Exhibit 1--continued.

Genuineness

A Scale for Measurement*

Level 5	All of the teacher's verbal communications are spontaneous. They appear to grow out of the current interaction only. They are not mechanical or practiced.
Level 4	Most of the teacher's verbal communications are spontaneous, but a few of them are ritualistic.
Level 3	The teacher's verbal communications are about equally distributed between ritualistic and spontaneous, but the spontaneous quality dominates slightly.
Level 2	Most of the teacher's verbal communications are mechanical, but a few of them are somewhat spontaneous.
Level 1	All of the teacher's verbal communications are ritualistic. They seem to be mechanical or practiced.

*David N. Aspy

Exhibit 1--continued.

Respect

A Scale for Measurement*

Level 5 The teacher consistently communicates a positive regard for the students' abilities to operate effectively at all intellective levels. Her guide is the students' direction. That is, she helps the students explore rather than directing them.

Level 4 The teacher consistently communicates a positive regard for the students' abilities to operate effectively in learning situations involving Level 1 of Bloom's Taxonomy, and occasionally allows her student to explore the higher intellective processes.

Level 3 The teacher consistently communicates a positive regard for the students' abilities to operate effectively in learning situations involving memory and recognition (Level 1 of Bloom's Taxonomy) but not with the higher intellective processes, i.e., creativity, problem solving, and evaluation.

Level 2 The teacher communicates a somewhat negative regard for the students' abilities to operate effectively in learning situations involving memory and recognition.

Level 1 The teacher communicates a clearly negative regard for the students' abilities to cope with any learning situation.

*David N. Aspy

Exhibit 2

FLORIDA CLIMATE AND CONTROL SYSTEM

INSTITUTE FOR THE DEVELOPMENT OF HUMAN RESOURCES
University of Florida
Gainesville, Florida
November 25, 1970

FLACCS*

Florida Climate and Control System
(1970-1971)

Program ____________________ Teacher ____________________
City ____________________ Date ____________________
School ____________________ Observer ____________________
Grade ____________________ Series ____________________

Children's Art Work Displayed

Abundant & varied	Quite a few	Some	A few	None
5	4	3	2	1

Relation of Room Displays and Artifacts To Children's Subcultural Background

Most are clearly related	Quite a few are related	Some are related	A few are related	None are related	Not applicable
5	4	3	2	1	0

*This is an experimental form which should not be cited or used without permission of the developers.

Exhibit 2--Continued

NEGATIVE AFFECT

C.	Tot	1	2	3	Verbal Teacher	C.	Tot	1	2	3	Nonverbal
10					Says "stop it," etc.	10					Waits for child
11					Uses threatening tone	11					Frowns
12					Rejects child	12					Points, shakes finger
13					Criticizes, blames	13					Pushes or pulls, holds
14					Warns	14					Shows disgust
15					Yells	15					Takes material
16					Scolds, humiliates	16					Refuses to respond to child
17					Other	17					Other
18					Code Involvement						

C.	Tot	1	2	3	Verbal Pupil	C.	Tot	1	2	3	Nonverbal
19					Says "No," "I won't" etc.	18					Makes face, frowns
20					Teases	19					Pouts, withdraws
21					Laughs	20					Uncooperative, resistant
22					Tattles	21					Stamps, throws, slams
23					Commands or demands	22					Interferes, threatens
24					Makes disparaging	23					Takes, damages property
25					Demands attention	24					Picks at child
26					Makes someone"feelsmall"	25					Pushes or pulls, holds
27					Finds fault	26					Hits, hurts
28					Threatens	27					Is left out
29					Other	28					Other
30					Code Involvement						

POSITIVE AFFECT

C.	Tot	1	2	3	Verbal Teacher	C.	Tot	1	2	3	Nonverbal
31					Says"Thank you,"etc.	29					Accepts favors for self
32					Agrees with child	30					Waits for child
33					Supports child	31					Gives individual attention
34					Gives indiv.attention	32					Warm, congenial
35					Warm, congenial	33					Listens carefully to child
36					Praises child	34					Smiles, laughs, nods
37					Develops"we feeling"	35					Pats, hugs, etc.
38					Is enthusiastic	36					Sympathetic
39					Other	37					Other
40					Code Involvement						

C.	Tot	1	2	3	Verbal Pupil	C.	Tot	1	2	3	Nonverbal
41					Says"Thank you,"etc.	38					Helpful, shares
42					Sounds friendly	39					Leans close to another
43					Agrees with another	40					Chooses another
44					Initiates contact	41					Smiles, laughs with another
45					Offers to share, co-operate	42					Pats, hugs another
46					Supports another	43					Agreeable, cooperative
47					Is enthusiastic	44					Enthusiastic
48					Praises another	45					Horseplay
49					Helps another	46					Other
50					Other	47					(Continued over)
51					Code Involvement						

CODE INVOLVEMENT

0. None involved
1. Few involved
2. Up to ½ the class
3. More than half

Exhibit 2--Continued

C.	Tot	1	2	3	TEACHER
10					Teacher Central
11					Leads singing, games, story tm
12					Moves freely among pupils
13					Withdraws from class
14					Uses blackboard, A-V Equip.
15					Ignores, refuses to attend P.
16					Attends P. briefly
17					Attends P. closely
18					Attends P. in succession
19					Attends simultaneous activ.
					VERBAL CONTROL
20					Praises
21					Asks for status
22					Suggests, guides
23					Feedback, cites reason
24					Questions for refltive. [illegible]
25					Correct w/o criticism (SM)
26					Questions for control
27					Questions state beh. rule
28					Directs with reason
29					Directs w/o reason
30					Uses time pressure
31					Call child by name (EWS)
32					Interrupts Pupil, cuts off
33					Warns
34					Supv. p. closely, imblizes.
35					Criticizes
36					Orders, commands
37					Scolds, punishes
38					Uses firm tone
39					Uses sharp tone
--					
--					
					NONVERBAL CONTROL
40					Tolerates deviant beh.
41					Positive redirection
42					Nods, smiles for control
43					Positive facial feedback
44					Uses "body English"
45					Gestures
46					Gives tangible reward
47					Touches, pats (gentle)
48					Holds, pushes, spanks (harsh)
49					Takes equipment, book
50					Signals, raps
51					Shhh! Shakes head
52					Glares, frowns

C.	Tot	1	2	3	PUPIL
10					Pupil Central
11					Pupil -- no choice
12					Pupil -- limited choice
13					Pupil -- free choice
14					(*Seat work w/o teacher
15					(*Seat work with teacher
16					(*Works, plays w. much supv.
17					(*Works, plays w. little supv.
18					(*Resists, disobeys directions
19					(*Obeys directions
20					Asks permission
21					Follows routine w/o reminder
22					Reports rule to another
23					Tattles
24					Gives information
25					Gives direction
26					Gives reason
27					Speaks aloud w/o permission
28					Engages in out-of-bounds beh.
29					Collaborates w. teacher
30					Task related movement
31					Aimless wandering
32					Fantasy
33					Uses play object as itself
34					Parallel play or work
35					Works, plays collaboratively
36					Works, plays competitively
37					Seeks reassurance, support
38					Shows pride
39					Shows fear, shame, humiliation
40					Shows apathy
					WORK GROUPS
41					Pupil as individual
42					Group w. teacher
43					Structured groups w/o t.
44					Free groups
					SOCIALIZATION
45					Almost never
46					Occasionally
47					Frequently
					MATERIALS
48					Structure T. behavior
49					Structure P. behavior
					PUPIL INTEREST ATTENTION
50-51					(Rank 1 low to 5 high)

Exhibit 3

SUMMARY OF CATEGORIES FOR THE RECIPROCAL CATEGORY SYSTEM

Category Number Assigned to Party 1*	Description of Verbal Behavior	Category Number Assigned to Party 2+
1	"WARMS" (INFORMALIZES) THE CLIMATE: Tends to open up and/or eliminate the tension of the situation; praises or encourages the action, behavior, comments, ideas and/or contributions of another; jokes that release tension not at the expense of others; accepts and clarifies the feeling tone of another in a friendly manner (feelings may be positive or negative; predicting or recalling the feelings of another are included).	11
2	ACCEPTS: Accepts the action, behavior, comments, ideas and/or contributions of another; positive reinforcement of these.	12
3	AMPLIFIES THE CONTRIBUTIONS OF ANOTHER: Asks for clarification of, builds on, and/or develops the action, behavior, comments, ideas and/or contributions of another.	13
4	ELICITS: Asks a question or requests information about the content subject, or procedure being considered with the intent that another should answer (respond).	14
5	RESPONDS: Gives direct answer or response to questions or requests for information that are initiated by another; includes answers to one's own questions.	15
6	INITIATES: Presents facts, information and/or opinion concerning the content, subject, or procedures being considered that are self-initiated; expresses one's own ideas; lectures (includes rhetorical questions -- not intended to be answered).	16
7	DIRECTS: Gives directions, instructions, orders and/or assignments to which another is expected to comply.	17
8	CORRECTS: Tells another that his answer or behavior is inappropriate or incorrect.	18
9	"COOLS" (FORMALIZES) THE CLIMATE: Makes statements intended to modify the behavior of another from an inappropriate to an appropriate pattern; may tend to create a certain amount of tension (i.e., bawling out someone, exercising authority in order to gain or maintain control of the situation, rejecting or criticizing the opinion or judgment of another).	19
10	SILENCE: Pauses, short periods of silence	
	CONFUSION: Periods of confusion in which communication cannot be understood.	20

---

*Category numbers assigned to Teacher Talk when used in classroom situation.

+Category numbers assigned to Student Talk when used in classroom situation.

APPENDIX II

Factor Structures

Exhibit 1

FACTOR STRUCTURE OF THE FLORIDA CLIMATE AND CONTROL SYSTEM

FLACCS Factor 1: Deviant Pupil Behavior vs. Teacher Response

Item	Loading	Description
11	.42	Teacher threatens, scolds w/consequence
14	.59	Teacher tolerates deviant behavior
31	.76	Pupil engages in out-of-bounds behavior
33	.72	Pupil aimless, wandering
42	.79	Frequent socialization by pupils
50	.46	Level 3 teacher nonverbal control
55	.78	Pupil teases
56	.65	Pupil commands or demands
57	.66	Pupil makes disparaging remarks
58	.70	Pupil threatens
70	.43	Pupil makes face
71	.69	Pupil uncooperative, resistant
72	.62	Pupil interferes, threatens
73	.56	Pupil takes, damages property of others
74	.80	Pupil picks at child
75	.78	Pupil pushes or pulls, holds
76	.76	Pupil hurts with something, hits
80	.48	Pupil leans close to another
81	.66	Pupil pats, hugs another
83	.82	Pupil horseplay
85	.79	Sum of pupil negative, verbal and nonverbal
27	-.69	Pupil obeys directions
41	-.59	Socialization by pupils
43	-.51	Pupil interest-attention

Exhibit 1--continued.

FLACCS Factor 2: Teacher Strong Control

Item	Loading	Description
1	.40	Teacher central
8	.61	Teacher questions for control
10	.63	Teacher directs, without reason
11	.61	Teacher threatens, scolds w/consequence
13	.68	Teacher gives orders, commands
17	.60	Teacher holds, pushes, punishes (other) isolates
18	.67	Teacher gestures, shhh, shakes head
19	.74	Teacher glares, frowns
25	.46	Pupil works, plays w/much supervision
31	.43	Pupil engages in out-of-bounds behavior
45	.50	Level 2 teacher verbal control
46	.64	Level 3 teacher verbal control
47	.71	Level 4 teacher verbal control
48	.75	Level 5 teacher verbal control
50	.71	Level 3 teacher nonverbal control
51	.79	Levels 4 & 5 teacher nonverbal control
52	.77	Teacher says "Stop it," etc.
53	.81	Teacher uses threatening tone
54	.63	Teacher criticizes
68	.78	Teacher frowns
69	.70	Teacher points, shakes finger
70	.50	Pupil makes face
71	.46	Pupil uncooperative, resistant
84	.85	Sum of teacher negative, verbal & nonverbal
85	.49	Sum of pupil negative, verbal & nonverbal
26	-.52	Pupil works, plays w/little supervision
27	-.42	Pupil obeys directions
43	-.50	Pupil interest-attention
60	-.51	Teacher sympathetic, warm

Exhibit 1--continued.

FLACCS Factor 3: Teacher Positive Behavior

Item	Loading	Description
5	.81	Teacher praises
6	.63	Teacher suggests
7	.64	Teacher questions for reflective thought
9	.64	Teacher directs, with reason
15	.55	Teacher nods, smiles for control
16	.56	Teacher touches, pats
30	.43	Pupil gives reason
36	.51	Pupil shows pride
38	.64	Group with teacher
44	.82	Level 1, teacher verbal control
45	.47	Level 2, teacher verbal control
49	.68	Levels 1 & 2, nonverbal control
61	.67	Teacher praises child
77	.56	Teacher waits for child (positive)
86	.51	Sum of teacher positive, verbal & nonverbal

FLACCS Factor 4: Pupil Free Choice vs. No Choice

Item	Loading	Description
21	.56	Pupil--limited choice
22	.60	Pupil--free choice
23	.71	Seat work without teacher
26	.52	Pupil works, plays w/little supervision
40	.46	Free groups
20	-.64	Pupil--no choice
24	-.67	Seat work with teacher
25	-.56	Works/plays with much supervision

Exhibit 1--continued.

FLACCS Factor 5: Pupil Supportive Behavior

Item	Loading	Description
28	.57	Pupil reports rule to another
30	.62	Pupil gives reason
34	.48	Pupil works, plays cooperatively, collaborates
35	.59	Pupil seeks reassurance, support
64	.47	Pupil agrees with another
65	.57	Pupil offers to share, cooperate
67	.70	Pupil helps another
79	.53	Pupil helpful, shares
87	.49	Sum of pupil positive verbal and nonverbal

FLACCS Factor 6: Teacher Positive Attention vs. Pupil Happy Response

Item	Loading	Description
3	.51	Teacher attends pupil briefly
4	.44	Teacher attends pupil closely
15	.56	Teacher nods, smiles for control
49	.52	Teacher sum of level 16.2 nonverbal teacher
59	.61	Teacher gives individual attention
60	.63	Teacher sympathetic, warm
62	.73	Teacher is enthusiastic, congenial
63	.51	Pupil sounds friendly
66	.68	Pupil is enthusiastic, happy
78	.69	Teacher smiles, laughs, nods
81	.57	Pupil pats, hugs another
86	.70	Sum of teacher positive verbal & nonverbal
87	.56	Sum of pupil positive verbal & nonverbal

Exhibit 2

FACTOR STRUCTURE OF THE RECIPROCAL CATEGORY SYSTEM

RCS Factor 1: Teacher Initiation vs. Pupil Response

Item	Loading	Description
2	.57	Teacher accepts. The sum of column 2.
4	.91	Teacher elicits. The sum of column 4.
12	.44	Pupil responds. The sum of 15.
20	.43	Student response to teacher. Rows 1-9 for column 15, divided by total student talk.
27	.91	Teacher narrow question. The sum of the 4-15 cell.
42	.54	Teacher elicit-initiate, percent. The column 4 total expressed as a percent of columns 4 plus 6.
44	.61	Teacher extended question. The 4-4 cell.
48	.51	Teacher talk. The sum of columns 1-9.
50	.89	Drill. The sum of the 4-15 cell plus the 15-4 cell.
61	.54	Total teacher talk (raw tallies) divided by number of batches.
10	-.42	Silence. The sum of column 10.
16	-.41	Confusion (does not include uncodeable tape). The sum of column 20.

Exhibit 2--continued.

RCA Factor 2: Student Talk vs.
Teacher Indirect/Direct Talk

Item	Loading	Description
12	.80	Pupil responds. The sum of 15.
21	.90	Pupil-pupil talk. The sum of rows 11-19 for columns 11-19.
47	.96	Steady-state student talk. The 11-19 diagonal.
49	.85	Student talk. The sum of columns 11-19.
52	.93	Inquiry. The sum of the 3-3 cell, plus the 4-4 cell, plus the 15-15 cell, plus the 16-16 cell.
53	.82	Inquiry-drill, percent. Inquiry divided by inquiry plus drill.
57	.51	Teacher indirect/direct, percent. The sum of column totals 1-3, divided by the sum of columns 1-3, plus 7-9.
17	-.76	Teacher talk, percent. The sum of cols. 1-9, divided by columns 1-9, plus columns 11-19.
20	-.52	Student response to teacher. Rows 1-9 for colum 15, divided by total student talk.

Exhibit 2--continued.

RCS Factor 3: Teacher Acceptance/Flexibility

Item	Loading	Description
1	.60	Teacher warms, informalizes the climate. The sum of column 1.
2	.60	Teacher acdepts. The sum of column 2.
3	.80	Teacher amplifies. The sum of column 3.
20	.40	Student response to teacher. Rows 1-9 for column 15, divided by total student talk.
22	.77	Teacher extended indirect. The sums of the cells in rows 1, 2, and 3 for columns 1, 2, and 3.
36	.72	Teacher-teacher flexibility. The number of cells in rows 1-9 for columns1-9 which are nonzer . (As percent of possible cells.)
40	.63	Total flexibility. The total number of cells in the entire 19 x 19 matrix (excluding row and column 10) which are nonzero. (As percent of possible cells.)
43	.69	Teacher amplify-direct, percent. The column 3 total expressed as a percent of column 3 plus column 7.
54	.68	Pupil initiation following teacher indirect. The sum of rows 1-3 for columns 11-19.
57	.51	Teacher indirect-direct, percent. The sum of column totals 1-3, divided by the sum of columns 1-3, plus 7-9.
61	.41	Total Teacher talk (raw tallies) divided by number of batches.

Exhibit 2--continued.

RCS Factor 4: Teacher Acceptance/Rejection

Item	Loading	Description
18	.91	Teacher acceptance-rejection, percent. The sum of columns 1, 2, and 3, divided by columns 1, 2, and 3, plus 8 and 9.
24	.74	Teacher revised I/D. This measure involved teacher indirect response (rows 11-19 for columns 1-3), and teacher direct response (rows 11-19 for columns 7-9). The percentage is made up of indirect response, divided by indirect response, plus direct response.
41	.89	Teacher accept-correct, percent. The column 2 total expressed as a percent of the sum of columns 2 and 8.
57	.57	Teacher indirect-direct, percent. The sum of column totals 1-3, divided by the sum of columns 1-3, plus 7-9.
8	.80	Teacher corrects. The sum of column 8.
9	.73	Teacher cools, formalizes. The sum of column 9.

Exhibit 2--continued.

RCS Factor 5: Extended Teacher Talk

Item	Loading	Description
6	.90	Teacher initiates. The sum of column 6.
17	.50	Teacher talk, percent. The sum of columns 1-9, divided by columns 1-9, plus columns 11-19.
45	.93	Steady-state teacher initiation. The 6-6 cell.
46	.90	Steady-state teacher talk. The 1-9 diagonal (that is, 1-1 plus 2-2 plus 3-3, etc.
48	.67	Teacher talk. The sum of columns 1-9.
51	.79	Average length of teacher initiation. The column 6 total divided by the column 6 total minus the 6-6 cell (calculated from the raw tally matrix).
58	.83	Teacher initiation, percent. The sum of column 6 divided by total teacher talk (columns 1-9).
10	-.47	Silence. The sum of column 10.
42	-.75	Teacher elicit-initiate, percent. The column 4 total expressed as a percent of columns 4 plus 6.

Exhibit 2--continued.

RCS Factor 6: Supportive Pupil Talk

Item	Loading	Description
25	.79	Pupil positive participation, percent. Positive participation divided by positive participation plus negative participation (rows 15, 16 for columns 11-13/rows 15, 16 for columns 11-13, plus rows 15, 16 for columns 17-19).
26	.83	Pupil revised I/D_1. This measure includes pupil extended indirect (rows 11-19 for columns 11-13); and pupil extended direct (rows 11-19 for columns 17-19); with pupil extended indirect divided by pupil extended direct, plus pupil extended indirect.
29	.45	Pupil broad question. The sum of the 14-16 cell.
30	.52	Pupil indirect interruption. The sum of row 6, column 11-13 cells.
39	.40	Pupil-pupil flexibility. The number of cells in rows 11-19 for columns 1-9 which are nonzero. (As percent of possible cells.)
59	.79	Pupil warms, accepts, amplifies. The sum of columns 11, 12, and 13.

Exhibit 2--continued.

RCS Factor 7: Teacher Direction/Criticism vs. Teacher Indirect Talk

Item	Loading	Description
7	.83	Teacher directs. The sum of column 7.
23	.89	Teacher extended direct. The sum of the cells in rows 7-9 for columns 7-9.
55	.90	Teacher direction and criticism. The sum of the following cells: 7-7, 7-9, 9-7, 9-9, 20-7, and 20-9.
31	-.44	Pupil substantive interruption. The row 6, columns 17-19 cells.
33	-.41	Total pupil interruption. The sum of pupil substantive interruption plus pupil direct interruption.
57	-.46	Teacher indirect-direct, percent. The sum of column totals 1-3 divided by the sum of columns 1-3 plus 7-9.

Exhibit 2--continued.

RCS Factor 8: Pupil Initiation vs. Teacher Response

Item	Loading	Description
5	.78	Teacher responds. The sum of column 5.
11	.84	Pupil elicits. The sum of 14.
31	.56	Pupil substantive interruption. The row c, columns 17-19 cells.
33	.58	Total pupil interruption. The sum of pupil substantive interruption plus pupil direct interruption.
35	.85	Pupil question, teacher response. The 14-5 cell.
37	.53	Teacher-pupil flexibility. The number of cells in the row 1-9 for columns 11-19, which are nonzero. (As percent of possible cells.)
38	.43	Pupil-teacher flexibility. The number of cells in rows 11-19 for column 1-9 which are nonzero. (As percent of possible cells.)

RCS Factor 9: Pupil Initiation vs. Pupil Response

Item	Loading	Description
13	.81	Pupil initiates. The sum of 16.
14	.53	Pupil directs. The sum of 17.
15	.47	Pupil corrects. The sum of 18.
19	.79	Pupil initiation percent. The sum of column 16 divided by total student talk. The sum of columns 11-19.
28	.56	Teacher broad question. The sum of the 4-16 cell.
38	.49	Pupil-teacher flexibility. The number of cells in rows 11-19 for column 1-9 which are nonzero. (As percent of possible cells.)
39	.59	Pupil-pupil flexibility. The number of cells in rows 11-19 for columns 11-19 which are nonzero. (As percent of possible cells.)

APPENDIX III

Regression Equations

Exhibit 1

REGRESSION EQUATIONS FROM ANALYSES USING HOLISTIC MEASURES AS PREDICTORS

Gain in

Anxiety = 19.76375 - 0.3733 PCMA* - 0.0255I.Q.
+ 1.0038SMNG + 5.3148SRES - 8.6976SGEN
+ 7.1574MMNG - 1.8266MRES - 8.8963MGEN
+ 2.4687MNG2 - 9.5676RXMG + 40.4454RES2
+ 2.0001GXMG - 57.4787GXRS + 22.4924GEN2.

Arithmetic = - 59.89598 - 0.4549PARH + 0.2846I.Q.
+ 0.8263SMNG - 9.3537SRES + 5.2517SGEN
- 3.8075MMNG + 1.8337MRES + 53.4683MGEN
- 21.1624MNG2 - 19.4329RXMG - 10.2572RES2
+ 65.3482GXMG + 38.2389GXRS - 63.8367GEN2.

Autonomy = 12.41003 - 0.7165PHS4 + 0.0493I.Q.
- 0.6726SMNG + 5.7857SRES - 4.9041SGEN
- 27.1046MMNG - 25.2206MRES + 50.5116MGEN
- 15.3458MNG2 + 58.4803RXMG - 15.8004RES2
- 8.4754GXMG - 27.8501GXRS + 9.0521GEN2.

Dependence-Proneness = - 10.03945 - 0.5520PD-P + 0.0345I.Q.
+ 0.4958SMNG - 6.2882SRES + 7.9844SGEN
- 16.6893MMNG - 59.5934MRES + 84.5555MGEN
+ 11.1316MNG2 - 78.8818RXMG - 9.7112RES2
+ 47.0184GXMG + 126.7415GXRS - 94.7911GEN2.

Flexibility = - 28.67097 - 0.8152PCFX + 0.2087I.Q.
+ 7.2902SMNG - 9.2296SRES + 0.0140SGEN
- 137.9856MMNG - 7.8423MRES + 195.0904MGEN
+ 39.2058MNG2 - 168.7049RXMG + 29.3000RES2
+ 114.0242GXMG + 136.9392GXRS - 159.2610GEN2.

Fluidity = - 54.02914 - 0.7293PCFU + 0.2081I.Q.
+ 10.9717SMNG - 8.9882SRES - 9.2056SGEN
- 203.2202MMNG + 22.0776MRES + 260.1123MGEN
+ 36.5968MNG2 - 110.8311RXMG + 7.9504RES2
+ 98.4877GXMG + 98.1084GXRS - 148.3433GEN2.

---

*Predictor terms beginning with P are pretest values of dependent variable.

Exhibit 1--continued.

Interpersonal Adequacy = 11.96601 - 0.7316PHS3 + 0.1097I.Q.
+ 6.5497SMNG + 5.9018SRES - 8.6333SGEN
- 97.3374MMNG - 51.1064MRES + 166.1366MGEN
- 10.4348MNG2 - 17.9016RXMG + 26.9649RES2
+ 68.8909GXMG - 6.6190GXRS - 63.0789GEN2.

JIM = - 30.06145 - 0.6351PJIM + 0.3546I.Q.
+ 3.9317SMNG - 49.1704SRES + 38.1771SGEN
+ 139.6017MMNG + 92.9717MRES - 174.6927MGEN
- 39.7414MNG2 - 255.0151RXMG
+ 95.4230RES2 + 241.4695GXMG + 92.0118GXRS
- 138.7946GEN2.

Lie Scale = 1.20564 - 0.7111PCML - 0.0172I.Q.
- 0.2414SMNG + 3.5073SRES - 1.3556SGEN
- 9.6900MMNG + 15.0967MRES + 0.0094MGEN
- 4.3945MNG2 + 24.6080RXMG - 9.2472RES2
- 6.8363GXMG - 17.2579GXRS + 10.7131GEN2.

Originality = - 56.75600 - 0.9033PCOR + 0.2446I.Q.
+ 6.1793SMNG - 11.9644SRES + 2.7550SGEN
- 213.9086MMNG - 0.6961MRES + 304.8115MGEN
- 32.3207MNG2 - 131.7537RXMG + 20.8608RES2
+ 269.4834GXMG + 110.5350GXRS
- 257.5120GEN2.

Physical Adequacy = 7.13478 - 0.6147PHS2 - 0.0078I.Q.
+ 6.2750SMNG - 3.0490SRES - 3.8627SGEN
- 7.6746MMNG - 64.6980MRES + 70.8697MGEN
- 4.5164MNG2 - 10.4951RXEP - 1.0241RES2
+ 17.3323GXMG + 39.5571GXRS - 38.0756GEN2.

Reading = - 30.42789 - 0.5395PREA + 0.3015I.Q.
- 2.9870SMNG + 3.0260SRES + 1.3485SGEN
+ 6.5506MMNG + 19.2686MRES + 5.8379MGEN
- 32.3379MNG2 + 52.6642RXMG - 0.0856RES2
+ 24.7455GXMG - 64.9024GXRS + 10.8434GEN2.

SRI = 20.35709 - 0.7264PSRI - 0.0475I.Q.
+ 0.4833SMNG + 3.9054SRES - 4.0512SGEN
+ 4.8365 MMNG + 14.3995MRES - 24.6734MGEN
- 6.5730MNG2 + 4.9484RXMG + 7.7456RES2
+ 7.9885GXMG - 24.0512 GXRS + 10.7380GEN2.

Exhibit 1--continued.

$$\begin{aligned}
\text{Spelling} ={}& -36.02144 - 0.4409\text{PSPL} + 0.3359\text{I.Q.}\\
&- 7.2637\text{SMNG} + 7.7625\text{SRES} + 1.8839\text{SGEN}\\
&- 60.7835\text{MMNG} - 91.7105\text{MRES} + 168.6808\text{MGEN}\\
&- 7.7669\text{MNG2} - 124.4622\text{RXMG} + 10.3549\text{RES2}\\
&+ 145.6322\text{GXMG} + 152.6422\text{GXRS}\\
&- 176.3495\text{GEN2}.
\end{aligned}$$

$$\begin{aligned}
\text{Teacher-School} ={}& -26.81380 - 0.6217\text{PHS1} + 0.0497\text{I.Q.}\\
&- 4.6431\text{SMNG} + 10.0918\text{SRES} - 0.2301\text{SGEN}\\
&- 46.1793\text{MMNG} - 63.7005\text{MRES} + 132.3121\text{MGEN}\\
&+ 11.8590\text{MNG2} + 24.0528\text{RXMG} - 13.6366\text{RES2}\\
&- 28.6639\text{GXMG} + 21.3391\text{GXRS} - 19.1211\text{GEN2}.
\end{aligned}$$

$$\begin{aligned}
\text{Vocabulary} ={}& 3.24302 - 0.5619\text{PVOC} + 0.3244\text{I.Q.}\\
&- 6.0677\text{SMNG} + 8.5412\text{SRES} - 0.7495\text{SGEN}\\
&+ 31.0760\text{MMNG} - 122.3900\text{MRES} + 68.2712\text{MGEN}\\
&- 16.0154\text{MNG2} + 17.3553\text{RXMG} - 43.6608\text{RES2}\\
&+ 6.4633\text{GXMG} + 102.9462\text{GXRS} - 60.0322\text{GEN2}.
\end{aligned}$$

Exhibit 2

REGRESSION EQUATIONS FROM ANALYSES USING SPECIFIC-BEHAVIOR MEASURES AS PREDICTORS

Gain on

Anxiety

$$\begin{aligned} &= 2.78935 - 0.4268\text{PCMA}^{*} - 0.0889\text{I.Q.}\\ &\quad + 0.0085\text{RCS1} + 0.0780\text{RCS2} - 0.0339\text{RCS3}\\ &\quad + 0.0641\text{RCS4} + 0.0184\text{RCS5} + 0.0102\text{RCS7}\\ &\quad + 0.0513\text{RCS8} + 0.0431\text{RCS9} + 0.0639\text{FLC2}\\ &\quad - 0.0252\text{FLC3} + 0.0214\text{FLC4} - 0.0324\text{FLC6}. \end{aligned}$$

Arithmetic

$$\begin{aligned} &= - 11.62547 - 0.5163\text{PARH} + 0.3713\text{I.Q.}\\ &\quad + 0.0768\text{RCS1} - 0.1062\text{RCS2} + 0.0930\text{RCS3}\\ &\quad - 0.2468\text{RCS4} - 0.0811\text{RCS5} + 0.2108\text{RCS7}\\ &\quad + 0.1310\text{RCS8} + 0.0404\text{RCS9} - 0.3164\text{FLC2}\\ &\quad + 0.5123\text{FLC3} + 0.1915\text{FLC5} - 0.4111\text{FLC6}. \end{aligned}$$

Autonomy

$$\begin{aligned} &= 5.20776 - 0.5436\text{PAUT} + 0.0561\text{I.Q.}\\ &\quad - 0.1012\text{RCS1} + 0.0389\text{RCS2} + 0.0621\text{RCS3}\\ &\quad + 0.0475\text{RCS4} - 0.0013\text{RCS5} + 0.0978\text{RCS7}\\ &\quad - 0.0208\text{RCS8} - 0.0870\text{RCS9} + 0.0398\text{FLC2}\\ &\quad + 0.1690\text{FLC3} - 0.0444\text{FLC5} - 0.0791\text{FLC6}. \end{aligned}$$

Dependence-Proneness

$$\begin{aligned} &= 19.70010 - 0.5686\text{PD-P} + 0.0455\text{I.Q.}\\ &\quad - 0.0546\text{RCS1} - 0.0758\text{RCS2} - 0.0257\text{RCS3}\\ &\quad - 0.0061\text{RCS4} + 0.0296\text{RCS5} + 0.0732\text{RCS7}\\ &\quad - 0.0119\text{RCS8} + 0.0097\text{RCS9} - 0.0915\text{FLC2}\\ &\quad + 0.0341\text{FLC3} + 0.0157\text{FLC5} - 0.0747\text{FLC6}. \end{aligned}$$

Flexibility

$$\begin{aligned} &= 17.09045 - 0.5822\text{PFLX} + 0.0038\text{I.Q.}\\ &\quad + 0.0569\text{RCS1} + 0.0157\text{RCS2} + 0.0323\text{RCS3}\\ &\quad - 0.0599\text{RCS4} - 0.0208\text{RCS5} + 0.0469\text{RCS7}\\ &\quad + 0.0732\text{RCS8} - 0.0290\text{RCS9} + 0.1044\text{FLC2}\\ &\quad + 0.0727\text{FLC3} + 0.0471\text{FLC5} - 0.0419\text{FLC6}. \end{aligned}$$

Fluidity

$$\begin{aligned} &= - 0.07922 - 0.4112\text{PFLU} + 0.0899\text{I.Q.}\\ &\quad + 0.0583\text{RCS1} + 0.0432\text{RCS2} + 0.0489\text{RCS3}\\ &\quad - 0.1017\text{RCS4} - 0.0029\text{RCS5} + 0.1239\text{RCS7}\\ &\quad + 0.1193\text{RCS8} + 0.0563\text{RCS9} + 0.1139\text{FLC2}\\ &\quad + 0.0191\text{FLC3} + 0.0729\text{FLC5} + 0.0056\text{FLC6}. \end{aligned}$$

---

*Predictor terms beginning with P are Pretest values of dependent variable.

Exhibit 2--continued.

Interpersonal
Adequacy = 58.56015 - 0.6976PINA + 0.1270I.Q.
- 0.0089RCS1 - 0.2421RCS2 + 0.1605RCS3
- 0.0172RCS4 - 0.0610RCS5 + 0.0144RCS7
- 0.3557RCS8 + 0.0517RCS9 - 0.0896FLC2
+ 0.2385FLC3 + 0.1025FLC5 - 0.2973FLC6.

JIM = 26.99216 - 0.5978PJIM + 0.3124I.Q.
- 0.0247RCS1 + 0.0079RCS2 + 0.1704RCS3
- 0.0294RCS4 + 0.2407RCS5 + 0.4927RCS7
- 0.1396RCS8 - 0.0648RCS9 - 0.7331FLC2
+ 0.2706FLC3 + 0.1767FLC5 - 0.3839FLC6.

Lie Scale = 0.87742 - 0.6083PLIE - 0.0037I.Q.
- 0.0215RCS1 - 0.0138RCS2 + 0.0341RCS3
+ 0.0233RCS4 + 0.0061RCS5 - 0.0207RCS7
- 0.0212RCS8 - 0.0225RCS9 + 0.0200FLC2
+ 0.0709FLC3 + 0.0180FLC5 - 0.0366FLC6.

Originality = - 4.91573 - 0.8221PORG + 0.1817I.Q.
- 0.0250RCS1 + 0.1238RCS2 + 0.2512RCS3
- 0.0639RCS4 - 0.0686RCS5 + 0.2822RCS7
- 0.0813RCS8 + 0.0970RCS9 - 0.1161FLC2
+ 0.1599FLC3 + 0.4825FLC5 - 0.4306FLC6.

Physical
Appearance = 25.85954 - 0.5123PPHA + 0.0139I.Q.
- 0.0443RCS1 - 0.1023RCS2 - 0.0454RCS3
- 0.0008RCS4 + 0.0057RCS5 + 0.0344RCS7
- 0.0189RCS8 + 0.0095RCS9 - 0.0822FLC2
+ 0.0900FLC3 + 0.0192FLC5 - 0.1294FLC6.

Reading = - 16.22656 - 0.4440PREA + 0.2541I.Q.
- 0.2140RCS1 + 0.2131RCS2 + 0.3413RCS3
- 0.0198RCS4 + 0.0793RCS5 + 0.1955RCS7
+ 0.0569RCS8 - 0.1586RCS9 - 0.1238FLC2
- 0.1166FLC3 - 0.0406FLC5 + 0.0684FLC6.

SRI = 17.80711 - 0.7610PSRI - 0.0520I.Q.
- 0.0092RCS1 + 0.0077RCS2 - 0.0173RCS3
- 0.0487RCS4 + 0.0001RCS5 - 0.0446RCS7
- 0.0088RCS8 + 0.0126RCS9 - 0.0111FLC2
- 0.0533FLC3 + 0.0279FLC5 + 0.0129FLC6.

Spelling = 11.67624 - 0.3937PSPL + 0.3767I.Q.
+ 0.4135RCS1 - 0.1656RCS2 + 0.0169RCS3
- 0.3936RCS4 - 0.0852RCS5 - 0.0392RCS7
- 0.1271RCS8 - 0.1358RCS9 + 0.3864FLC2
+ 0.1934FLC3 - 0.4028FLC5 - 0.4065FLC6.

Exhibit 2--continued.

$$\begin{aligned}\text{Teacher-School} = {} & 16.41690 - 0.6272\text{PTSL} + 0.0519\text{I.Q.}\\ & - 0.0876\text{RCS1} - 0.1260\text{RCS2} + 0.0133\text{RCS3}\\ & + 0.0339\text{RCS4} + 0.1372\text{RCS5} + 0.0426\text{RCS7}\\ & - 0.0889\text{RCS8} + 0.0123\text{RCS9} - 0.0939\text{FLC2}\\ & + 0.1111\text{FLC3} - 0.0258\text{FLC5} - 0.0838\text{FLC6}.\end{aligned}$$

$$\begin{aligned}\text{Vocabulary} = {} & 1.33651 - 0.5502\text{PVOC} + 0.3676\text{I.Q.}\\ & - 0.0030\text{RCS1} + 0.0385\text{RCS2} + 0.1293\text{RCS3}\\ & - 0.2633\text{RCS4} + 0.1615\text{RCS5} - 0.0105\text{RCS7}\\ & - 0.2029\text{RCS8} + 0.0996\text{RCS9} - 0.2649\text{FLC2}\\ & + 0.1096\text{FLC3} + 0.1937\text{FLC5} - 0.1267\text{FLC6}.\end{aligned}$$

APPENDIX IV

Items from Scoring of Categorical Instruments

Exhibit 1

ITEMS SCORED FOR FLORIDA CLIMATE AND CONTROL SYSTEM

Teacher

1 Teacher central
2 Ignores pupil
3 Attends pupil briefly
4 Attends pupil closely

Teacher Verbal Control

5 Praises
6 Suggests
7 Questions for reflection, thought
8 Questions for control
9 Directs, with reason
10 Directs, without reason
11 Threatens, scolds w/consequence
12 Supervises pupil closely
13 Gives orders, commands

Teacher Nonverbal Control

14 Tolerates deviant behavior
15 Nods, smiles for control
16 Touches, pats
17 Holds, pushes, punishes, (other), isolates
18 Gesture, shhh, shakes head
19 Glares, frowns
20 Pupil--no choice
21 Pupil--limited choice
22 Pupil--free choice
23 (Seat work w/o teacher*)
24 (Seat work with teacher*)

Teacher

25 (Works, plays w/much supervision*)
26 (Works, plays w/little supervision*)
27 (Obeys directions*)
28 Reports rule to another
29 Gives direction
30 Gives reason
31 Engages in outof-bounds behavior
32 Task-related movement
33 Aimless wandering
34 Works, plays cooperatively, collaborates
35 Seeks reassurance, support
36 Shows pride

Work Groups

37 Pupil as individual
38 Group with teacher
39 Structured groups
40 Free groups

Socialization

41 Never
42 Frequently
43 Fifth grade pupil interest attention
44 Sum of level 1 verbal
45 Sum of level 2 verbal
46 Sum of level 3 verbal
47 Sum of level 4 verbal
48 Sum of level 5 verbal
49 Sum of level 1 & 2 nonverbal
50 Sum of levels 4 & 5 nonverbal

Exhibit 1--continued.

Teacher Negative Verbal

52 Says "Stop it," etc.
53 Uses threatening tone
54 Criticizes

Teacher Negative Nonverbal

68 Waits for child
69 Points, shakes finger

Pupil Negative Verbal

55 Teases
56 Commands or demands
57 Makes disparaging remarks
58 Threatens

Pupil Negative Nonverbal

70 Makes face
71 Uncooperative, resistant
72 Interferes, threatens
73 Takes, damages property of others
74 Picks at child
75 Pushes or pulls, holds
76 Hurts with something, hits

Teacher Positive Verbal

59 Gives individual attention
60 Sympathetic, warm
61 Praises child
62 Is enthusiastic, congenial

Teacher Positive Nonverbal

77 Waits for child
78 Smiles, laughs, nods

Pupil Positive Verbal

63 Sounds friendly
64 Agrees with another
65 Offers to share, cooperate
66. Is enthusiastic, happy
67 Helps another

Pupil Positive Nonverbal

79 Helpful, shares
80 Leans close to another
81 Pats, hugs another
82 Enthusiastic, happy
83 Horseplay
84 Sum of teacher negative verbal and nonverbal
85 Sum of pupil negative verbal and nonverbal
86 Sum of teacher positive verbal and nonverbal
87 Sum of pupil positive verbal and nonverbal

Exhibit 2

VARIABLES FOR RECIPROCAL CATEGORY SYSTEM

Variable	Description
1	Teacher warms, informalizes the climate. The sum of column 1.
2	Teacher accepts. The sum of column 2.
3	Teacher amplifies. The sum of column 3.
4	Teacher elicits. The sum of column 4.
5	Teacher responds. The sum of column 5.
6	Teacher initiates. The sum of column 6.
7	Teacher directs. The sum of column 7.
8	Teacher corrects. The sum of column 8.
9	Teacher cools, formalizes. The sum of column 9.
10	Silence. The sum of column 10.
11	Pupil elicits. The sum of 14.
12	Pupil responds. The sum of 15.
13	Pupil initiates. The sum of 16.
14	Pupil directs. The sum of 17.
15	Pupil corrects. The sum of 18.
16	Confusion (does not include uncodeable tape). The sum of column 20.
17	Teacher talk, percent. The sum of columns 1-9 divided by columns 1-9 plus columns 11-19.
18	Teacher acceptance-rejection, percent. The sum of columns 1, 2, and 3 divided by columns 1, 2, and 3 plus 8 and 9.

Exhibit 2--continued.

Variable	Description
19	Pupil initiation. The sum of column 16 divided by total student talk. The sum of columns 11-19.
20	Student response to teacher. Rows 1-9 for column 15, divided by total student talk.
21	Pupil-pupil talk. The sum of rows 11-19 for columns 11-19.
22	Teacher extended indirect. The sums of the cells in rows 1, 2, and 3 for columns 1, 2, and 3.
23	Teacher extended direct. The sum of the cells in rows 7-9 for columns 7-9.
24	Teacher revised I/D. This measure involves teacher indirect response (rows 11-19 for columns 1-3), and teacher direct response (rows 11-19 for columns 7-9). The percentage is made up of indirect response divided by indirect response plus direct response.
25	Pupil positive participation, percent. Positive participation divided by positive participation plus negative participation (rows 15, 16 for columns 11-13/rows 15, 16 for columns 11-13, plus rows 15, 16, for columns 17-19).
26	Pupil revised I/D$_1$. This measure includes pupil extended indirect (rows 11-19 for columns 17-19), with pupil extended indirect divided by pupil extended direct, plus pupil extended indirect.
27	Teacher narrow question. The sum of the 4-15 cell.
28	Teacher broad question. The sum of the 4-16 cell.
29	Pupil broad question. The sum of the 14-16 cell.
30	Pupil indirect interruption. The sum of row 6, column 11-13 cells.
31	Pupil substantive interruption. The row 6, columns 17-19 cells.

Exhibit 2--continued.

Variable	Description
32	Pupil direct interruption. The row 6, columns 17-19 cells.
33	Total pupil interruption. The sum of pupil substantive interruption plus pupil direct interruption.
34	Pupil question, teacher question. The 14-14 cell.
35	Pupil question, teacher response. The 14-5 cell.
36	Teacher-teacher flexibility. The number of cells in rows 1-9 for column 1-9 which are nonzero. (As percent of possible cells.)
37	Teacher-pupil flexibility. The number of cells in the rows 1-9 for columns 11-19, which are nonzero. (As percent of possible cells.)
38	Pupil-teacher flexibility. The number of cells in rows 11-19 for column 1-9 which are nonzero. (As percent of possible cells.)
39	Pupil-pupil flexibility. The number of cells in rows 11-19 for columns 11-19 which are nonzero. (As percent of possible cells.)
40	Total flexibility. The total number of cells in the entire 19 x 19 matrix (excluding row and column 10) which are nonzero. (As percent of possible cells.)
41	Teacher accept-correct, percent. The column 2 total expressed as a percent of the sum of columns 2 and 8.
42	Teacher elicit-initiate, percent. The column 4 total expressed as a percent of columns 4 plus 6.
43	Teacher amplify-direct, percent. The column 3 total expressed as a percent of column 3 plus column 7.
44	Teacher extended question. The 444 cell.

Exhibit 2--continued.

Variable	Description
45	Steady-state teacher initiation. The 6-6 cell.
48	Teacher talk. The sum of columns 1-9.
49	Student talk. The sum of columns 11-19.
50	Drill. The sum of the 4-15 cell plus the 15-4 cell.
51	Average length of teacher initiation. The column 6 total divided by the column 6 total minus the 6-6 cell (calculated from the raw tally matrix).
52	Inquiry. The sum of the 3-3 cell plus the 4-4 cell plus the 15-15 cell plus the 16-16 cell.
53	Inquiry-drill, percent. Inquiry divided by inquiry plus drill.
54	Pupil initiation following teacher indirect. The sum of rows 1-3 for columns 11-19.
55	Teacher direction and criticism. The sum of the following cells: 7-7, 7-9, 9-7, 9-9, 20-7, and 20-9.
56	Pupil direction and criticism. The sum of the following cells: 17-17, 17-19, 19-17, 19-19, 20-17, and 20-19.
57	Teacher indirect-direct, percent. The sum of column totals 1-3 divided by the sum of columns 1-3 plus 7-9.
58	Teacher initiation, percent. The sum of column 6 divided by total teacher talk)columns 1-9).
59	Pupil warms, accepts, amplifies. The sum of columns 11, 12, and 13.
60	Total number of tallies for all batches (raw).
61	Total teacher talk (raw tallies) divided by number of batches.

BIOGRAPHICAL SKETCH

Flora Nell Roebuck is a native of New Bern, North Carolina. She is a graduate of St. Paul's High School, St. Paul, N. C. She attended Wake Forest College in Winston-Salem, N.C. and graduated from the University of North Carolina at Chapel Hill, with the B. A. in radio television and motion pictures and Honors in Creative Writing. As a Ford Foundation Fellowship recipient, she earned the degree of Master of Education from Cornell University, Ithaca, New York. She is a member of Pi Lambda Theta and of several professional organizations. She has held public school administrative or teaching positions in four states. She was an instructor in the Foundations Department of the College of Education at the University of Florida and assistant professor of the Curriculum and Instruction Department of the College of Education of Northeast Louisiana University. She is currently directing the development of the Texas Right to Read training modules in a cooperative program between the Texas Education Agency and Texas Woman's University. She is the originating author and consulting editor of the Random House Individualized Reading Program.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Education.

Donald L. Avila

Donald L. Avila
Professor of Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Education.

Arthur W. Combs

Arthur W. Combs
Professor of Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Education.

Walter B. Busby

Walter B. Busby
Associate Professor of Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Education.

Robert B. Primack
Associate Professor of Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Education.

Franz R. Epting
Associate Professor of Psychology

This dissertation was submitted to the Graduate Faculty of the College of Education and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Education.

August, 1975

B. J. Sharp by M C Baker
Dean, College of Education

Dean, Graduate School